2nd Annual Visual and Iconic Language Conference

July 21st & 22nd, 2008
San Diego, California

Space and Naval Warfare Systems Center, Pacific

SPAWAR
Systems Center

Keynote Speakers

David Gray
CEO & Founder
XPLANE

Visual literacy: Toward a visual alphabet

Humans learn language in both natural and formal ways. In natural languages such as English, children learn to speak without formal instruction, while reading and writing must be methodically taught. In the absence of formal instruction, children typically do not learn to read and write. Formal languages, such as mathematics and computer programming must also be taught, because they are based on an explicit set of rules that do not emerge naturally. Language teaching begins with the introduction of basic elements, followed by rules that are used to combine the elements into more complex patterns. Letters are introduced first. Letters are then used to construct words, words are used to construct sentences, sentences are used to construct paragraphs, and so on. For example, instruction in English begins with the ABC book, where the alphabet is introduced by connecting symbols with sounds, as well as pictures of tangible objects. For example, A is for Apple, B is for Ball. After mastering the alphabet, a student moves on to construct words and then simple sentences. The same method is used to teach mathematics, which begins with the introduction of numbers and counting, followed by basic equations like 2+2=4, after which more complex patterns can be considered. This phased approach to learning allows students to grasp the most basic concepts and structures before combining them into complex patterns. Visual language, if it is a language, has no shared framework for formal instruction. A visual alphabet, vocabulary and grammar have yet to be defined and standardized. If literacy is defined as the ability to read and write a language, then a visually literate society requires a shared visual alphabet and grammar. In this talk, a visual alphabet and some basic rules of construction are proposed. The concepts are a synthesis of academic writings on visual language and observations of how visual language functions naturally in human interaction.

Alan Stillman
CEO & Founder
Kwikpoint

Practical Applications of Visual Language

Experiences from the field in applying Kwikpoint visual language translators to real-world problems in military operations.

Neil Cohn
Psychology PhD Candidate,
Tufts University

What is "Visual Language"

Many theories describing "visual language" have been emerging from diverse fields including computer science, communications, and design. However, often these approaches rely on metaphoric or folk notions of "language" without delving deeper into what language actually consists of, especially on a cognitive level. This talk will present Visual Language Theory from the view of linguistics and cognitive science to discuss what "language" entails, and thereby exploring just what it means to have a literal theory of a graphic modality of language. The result will be a view of graphic communication and the capacity for drawing that is embedded alongside other mental capacities and divorced from socio-cultural labels that stymie its recognition.

Conference Background

In the fall of 2006 SPAWAR Systems Center began a new program to assess and augment tactical and humanitarian communications capabilities using visual & iconic language. The US DoD, NATO, Department of Homeland Security, and many others have each separately devoted substantial effort to the creation of standard warfighting, hazard, and security symbols and glyphs for use within operations visualization systems (such as the Standard Warfighting Symbology, MIL STD 2525). These standard symbols are used to represent various unit and hazard types, attributes, and states within common operational picture displays. The symbols within these standards are, however, often ill-suited to efficient and clear communication in a tactical setting. They appear sufficient for many mapping systems, providing detailed information within the context of a terrain or battlefield, but are generally not easily used as communicable entities or sequences within messages which may be primarily textual. This and other issues produce a communications lapse between individuals which may or may not have access to the same view of the battlespace or humanitarian operation.

Our primary research goal is to utilize visual language (to include gesture and sign) to produce integrative, multi-modal technologies to address communications shortfalls latent within modern communications techniques in use by the DoD and others, including; email, chat, text-messaging, DMS, web-based (e.g. wiki, blog, etc), video, audio, and even direct line-of-sight communications. Scientific methods for the creation and assessment of visual languages and visual communication strategies are of particular interest to us, including approaches in computational linguistics, cognitive science, and theory developed from the study of the historical evolution of language both natural and constructed.

Overall, the VAIL conference is intended to foster an open forum, allowing visual language researchers to gather with other experts and share research findings and hold open discussions. We had a very successful first year and expect the conference and community to grow during the next several years.

Within the context of the VaIL conference it is intended for the term 'visual language' to refer to all visual perceptual communication techniques, including, but not limited to: representational and non-representational glyphs and icons, graphic design, signed language, gesture, acronyms, and visual aspects of all written language (ideographic, pictographic, syllabaries, or alphabets.)

Conference Agenda

Day 1 - Monday, July 21st

07:30-08:30 Registration
09:00-09:20 Introduction
09:20-10:20 **Dave Gray**
Visual literacy: Toward a visual alphabet
XPlane

10:25-11:05 **Marion Ceruti**
Application of Maya Hieroglyphic Principles to Military Symbology and Communication
SPAWAR Systems Center, Pacific

11:10-11:50 **Geoffrey Draper**
A Visual Query Language for Correlation Discovery and Management
University of Utah

11:50-12:50 Lunch
12:50-13:30 **Juan Wachs**
Considerations on Optimal Hand Gestures Design
Naval Postgraduate School

13:35-14:15 **Valerie Sutton & Adam Frost**
SignWriting: Sign Languages Are Written Languages
Center for Sutton Movement Writing

14:15-14:25 Break
14:25-15:05 **Alan Stillman**
Applications of visual language
Kwikpoint

15:10-16:30 Panel on
Visual language production, navigation, and comprehension

Conference Agenda

Day 2 - Tuesday, July 22nd

09:00-09:40 LorRaine Duffy
Tactical Situation Assessment Technologies
SPAWAR Systems Center, Pacific

09:45-10:25 Vincent Dinh
A non-optical method for capturing the muscular-skeletal configuration of a human body part
SPAWAR Systems Center, Pacific

10:30-11:30 Neil Cohn
What is "Visual Language"
Tufts University

11:30-12:40 Lunch
12:40-13:40 Panel on Gesture writing systems
13:45-14:40 Panel on Cross-cultural and cross-domain visual communication
14:40-15:00 Closing Remarks

SPAWAR Pacific Conference Organizers:

Sunny Fugate Conference Chair fugate@spawar.navy.mil	Marion Ceruti, PhD Program Chair marion.ceruti@navy.mil	Emily Medina Local Arrangements Chair ewilson@spawar.navy.mil

The VaIL 2008 conference is organized and hosted by the Space and Naval Warfare Systems Center San Diego, Pacific and sponsored by the Office of Naval Research through a Science and Technology 'Grand Challenge' research program. We would like to thank the Space and Naval Warfare Systems Center for the opportunity to hold the VaIL conference and the Office of Naval Research for providing much sought after funding for basic research.

Application of Maya Hieroglyphic Principles to Military Symbology and Communication

Marion G. Ceruti, Ph.D., Sunny Fugate, Jeffrey Ellen,
Emily Medina, LorRaine Duffy, Ph.D., and Vincent V. Dinh, Ph.D.
Space and Naval Warfare Systems Center, San Diego (SSC-SD)
53560 Hull Street, San Diego, CA 92152-5001
Tel. 619 553-4068, FAX 619 553-3931, marion.ceruti@navy.mil, lorraine.duffy@navy.mil

Abstract - *The Mayan languages have been spoken for over 5,000 years. The language structure, syntax, and semantics of Mayan languages are rich and varied. This paper describes how to apply some principles of the Maya Hieroglyphic language to interactive command-and-control displays in decision-support system. It also suggests some preliminary elements of a tactical-assessment language where message components, including input transmitted from hand gestures, can be expressed as icons and integrated into tactical chat for situation-assessment reports. A large part of the Mayan language is expressed in terms of facial glyphs; faces are part of any human experience, regardless of culture. This suggests that a relationship between facial depictions and fine structure in glyphs and affixes may be easier for users to remember than a randomly or even a sequentially ordered array of glyphs and affixes. The paper concludes with directions for future research.*

Keywords- Iconic language, glyph, Maya language, Maya hieroglyphics, affix, symbol, data glove, chat

1. Introduction

Military use of a Native North American language, or some aspect thereof, is not new. During World War II, Navajo native speakers, called "code talkers," used a code based on their language to send secure military messages in radio communications [10].

This paper reports the findings of a study of the structural principles of Maya hieroglyphic language that could be applied to military communications. The written Maya hieroglyphic script is logographic, similar to other well-known hieroglyphic scripts, but is augmented with a set of syllabic glyphs used to specify pronunciation [5], [9]. Thus, the Maya script is a logosyllabic system [5]. Maya hieroglyphic script is well organized, logical, and can be learned quickly [5], especially the parts of the language that pertain to the language structure.

In addition to these benefits, the Maya hieroglyphic language is of interest because of its parallel to mathematical groups. The language's use of affixes suggests well-defined combining operations. In any system, the ability to represent the system elements uniquely is important. The Maya language is context dependent, which suggests relativity and that the basis symbol set used is a maximal one (eigen-set). This sort of representation characteristic is effective for any arbitrary communication class. Also, the language's representation style is reader-interpretation independent, which promotes message clarity and uniqueness.

Research into pictorial and iconic languages is part of a larger project called Tactical Situation Assessment Technology (TSAT). The objective of the TSAT program is to produce a) the next generation in tactical chat, b) icon-based situation assessment (SA) language and c) wireless gesture-recognition and transmission gloves (called "data gloves") for use in hostile and other extreme environments. The three main research thrusts in TSAT will become integrated in more advanced research projects of the future. The present study supports the development of a prototype iconic language for SA in command and control (C^2). (See, for example, [2].)

2. Context

Webster [8] defines an icon as a pictorial representation or an image. This definition implies that an icon is context free, but this is too broad to be useful in a C^2 environment. An icon also has been defined [7] as a presentational element that designates an object of reference and specifies its "locus" in the presentation context. An icon in this usage does not qualify or describe the entity it signifies but acts as a pointer [7].

Icons are completely context dependent when used as arbitrary pointers. Such icons lose all meaning if context is dropped. Representational icons directly or indirectly describe their referents; therefore, they are iconic to the referent. Iconic forms are often context sensitive, but by definition, they are information laden even when context is removed.

Context-independent icons are not constrained to denoting objects, i.e. nouns. Other parts of speech, such as adjectives, can be represented efficiently using a grammar internal to the icon. The important part of the definition that is relevant to this discussion is the idea that the icon and its context convey information.

As is the case with all languages, the context of complete thoughts, expressed in the form of complete sentences, facilitate a comprehensive understanding of an iconic language. Thus, a visual, iconic language consists of concatenated icons that form sentences [6] much in the same way that concatenated characters, syllables, and words form complete sentences in a phonetic language. With any language, the method of concatenation depends on the structure and grammar rules of the language. The design of a good iconic or symbolic language will be based on an ontology and knowledge representation that takes advantage of the two-dimensional flexibility and potentially intuitive nature of icons. (See, for example, [3] and [6].) The factors that contribute to the context of an iconic or pictorial language include but are not limited to, many of the factors that contribute to context in spoken, phonetic languages.

Ideally, to simplify design, clarify communications, and facilitate comprehension, we often attempt to evolve the usage of a context-independent grammar with respect to the design and use of any icon, glyph, or symbol. This is a design aspect of MIL STD 2525B [4] in which the icons can be understood in a context-less environment but their relationships and importance are conveyed solely by their presentation context. This is sensible, as the nature of the battle space, its graphical depictions (e.g. displays designed to communicate the Common Operating Picture (COP)) and text written about specific elements or events in the battle space preclude any comprehensive understanding of elements or events surrounding them in the absence of context.

Context, including information pedigree [1], is very important in situational assessment, threat assessment, and resource allocation. The relationships, intentions, plans, capabilities, and the feedback loops [11] between objects in the battle space are elements that constitute the context. Course-Of-Action (COA) generation relies on timely context discovery and the reduction of uncertainty with respect to context.

3. Glyphs

A *glyph* is a structured representation coded to portray relevant semantics of the associated object or group of objects [7]. It is basically an icon with modifiers that provides access to information about the entity denoted on its face [7]. The information may be data about the entity or about available redirection to another presentational device where additional information may be obtained [7]. These definitions of glyph relate remarkably well to the notion of glyph described in [4].

In the Mayan script, a Maya *glyph block* is a group or aggregate of symbols taken as a unit and occupying a cell in a grid of depicted symbols [5]. These grids consist of glyph blocks arranged in columns and rows similar to a table [5]. Maya glyph blocks have two or more signs and usually consist of a *main sign* or icon and associated affixes, or modifiers of the main sign [5]. Sometimes a glyph block can consist of two main signs. Maya icons are distinguished by their size and position within the glyph block. However, in some cases main signs also can serve as modifiers, indicating that size is not the only determinant of a sign's function, whereas affixes also can act as syllables and pronunciation clues [5].

The main sign represents a whole word (as opposed to a part of a word, such as a letter of a syllable) and the glyph block forms the main aggregate in the Maya writing system [5]. It can be conceptualized as words in a sentence [5]. Here, as in C^2, use depends on context because the meaning of many glyphs forms the context in which any single glyph must be understood.

4. Affixes

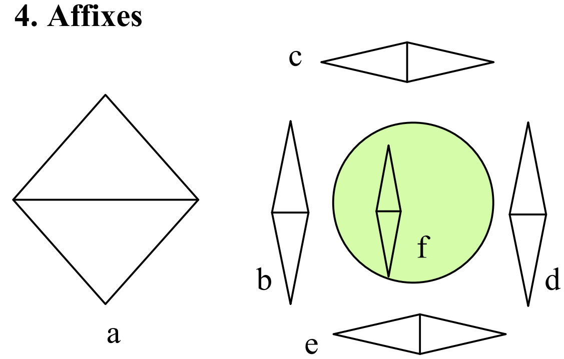


Figure 1. Glyph as a) main sign, or as extraglyphic b) prefix, c) superfix, d) postfix, e) subfix, and intraglyphic f) infix.

Affixes consist of infixes, prefixes, superfixes, postfixes and subfixes [5] depending on where in the glyph block they occur. Prefixes, superfixes, postfixes and subfixes can be intraghyphic or extraglyphic depending on how the display context is defined. Figure 1 depicts how an icon or glyph can be used as either a main sign or an affix. Infixes, as the name implies, are always intraglyphic. (See, for example, [5]). Affixes can be used as main signs, words, syllabic elements, phonetic complements, and semantic determinatives [4].

5. Applications

5.1 Interactive displays in decision-support systems

The MIL STD 2525B [3] symbology was designed for use with geographical maps [5]. MIL STD 2525B symbology is limited to coarse-grained depiction of single elements with limited semantics, detail, and granularity [5]. The use of affixes based on the Maya system can enhance and refine the semantics and granularity of MIL STD 2525B without significantly increasing the complexity of the symbols. How to do this in an intuitive way that requires minimal training with respect to operator comprehension remains an open research question.

We need to depict various kinds and levels of uncertainty in glyphs and icons that report states of the battle space [7]. The Maya-glyph structure and position of affixes (e.g. infixes, prefixes, superfixes, postfixes and subfixes [5] as depicted in Figure 1) provide an efficient model for future military-glyph design.

One possible application is that affixes could be used to construct and utilize *knowledge glyphs* [7], which enable the user to access from the glyph, extraglyphic information that pertains to the entity that the glyph represents. Knowledge glyphs should be internally coherent, referentially comprehensive and explicitly linked to visual interface elements most important to a commander in selecting a COA [7].

For example, if a MIL STD 2525B symbol representing a surface ship were augmented with affixes, each affix could be assigned a different data set that points to amplifying information about the ship. The assignment of data sets to various positions in the "glyph block" could be systematized and standardized so that the user would know which affix to select to display a particular data set.

The user could select a data set by clicking on the affix and the pop-up window could display the data on the screen, or on the original display. Thus, an array of affixes and a main-sign glyph could be used to denote and to facilitate access to the merged context of multiple-characteristic data sets, thus providing much more fine structure than the nature of the glyph or symbol alone would suggest. This application of affixes extends the work described in [7] by suggesting a systematic approach to the construction and usage of knowledge glyphs.

To merge the concepts that the terminology in [5] and [7] represent, a knowledge glyph together with its affixes that act as pointers to knowledge or additional information can be termed *knowledge glyph block*. The advantage of using knowledge glyph blocks over the more common "drill-down" features of "point-and-click" on a single symbol is that it enables the user to select specific information categories individually. The user does not see all the information available about the entity the glyph represents, just a smaller topic-oriented data set. Knowledge glyphs serve as access points or intersections for two or more referential contents [7].

5.2 Phonetic affixes and symbolic composition

A perplexing result of modern information design is a dearth of phonetic information. For the most part, military symbology cannot be verbalized. The symbol must be explained using detailed descriptions instead of being encapsulated by a single verbalization. This limits the communication media that can be used efficiently to communicate the idea(s) represented by the glyph. A common resolution to this problem is to use acronyms to provide a convenient pointer to the concept. However, an acronym, like the concept that it represents, often is not unambiguously pronounceable. Too many of the verbal cues have been stripped. Many examples have been given consistent pronunciation (e.g. GUI or SQL, EW), but the correct pronunciation is still ambiguous.

In the Maya hieroglyph script are both logograms and syllabograms that explicitly denote the verbal representation for the concept. This directly ties together the visual and the verbal representations without compromising the visual representation.

The use of phonetic affixes could be adapted for military symbology to provide consistent, unambiguous verbalizations for iconography. The verbal forms need not to rely on existing words. However, sharing common phonemes would assist both in the acquisition and use of coded verbal forms. As in the Maya script, the pronunciation codes could themselves be hieroglyphs or even pictographic [3]. The potential benefits in using this approach for tactical communications cannot be overstated.

Modern-day acronyms are words that have been stripped out of their linguistic context leaving only a sequence of arbitrary symbols. This is essentially a hieroglyphic form that retains phonetic cues. However, due to the compactness of acronyms they are ambiguous across domains and social circles and thus context dependent. (For example, SME means “subject-matter expert” or “small-to-medium enterprise” depending on context.)

We should learn from Mayan cultures in both regards for symbol-pronunciation ambiguity and rules for language creation. That we assign new acronyms so easily indicates the need to streamline communications by compressing long names of entities to which we must refer often. The frequent creation and use of acronyms also could indicate a societal need for new words without waiting for widespread usage and without the disadvantages of formality. Acronyms are essentially abbreviations.

This likely indicates that we are too restrictive in our cultural and philosophical policies toward emergent language. And, seemingly contradictorily, we need to loosen our grasp on language to foster its evolution. Languages evolve in many ways. Restricting or mandating a particular language usage through rules can be successful only among groups over which a leader can exercise control.

Assuming such control can be exercised, in essence, this means providing ‘word-creation’ rules for the layman. For example, one such rule might be that acronyms must have sufficient phonetic cues. This could be accomplished by selecting one among the following multiple alternatives:

a) The traditional period, ‘.’, between characters can be used to indicate not only abbreviation, but also to indicate each character needs to be pronounced individually (e.g. ‘G.U.I.’, ‘U.C.L.A’, ‘D.O.D’)
b) Explicit phoneticization could be implemented (e.g. GooUI vs. “GUI,” SeeQueL vs. “SQL,” and “VerSus” vs. “vs.”)
c) Select names for new entities that will result in acronyms that can be remembered easily and also can be pronounced as a word without the necessity to add extra characters.

Examples of c) are as follow: Light Amplification by Stimulated Emission of Radiation (LASER), and Self-Contained Breathing Apparatus (SCUBA), and SOund NAvigation Ranging (SONAR). Clearly, alternative c) is preferable to a) and b) if and when possible. After all, the purpose of using the acronym in the first place is to avoid the use of a long name.

5.3 Facial depictions and fine structure in glyphs and affixes

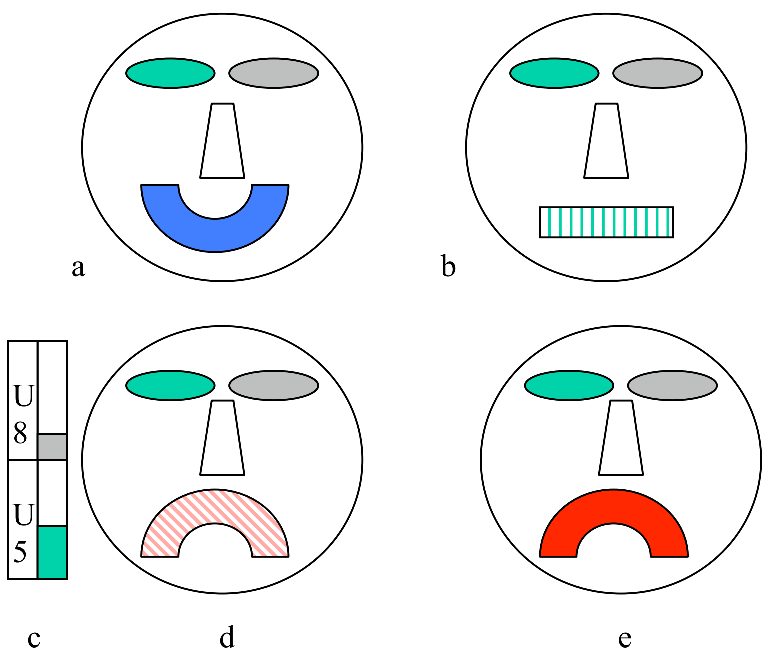


Figure 2. Four main-sign facial glyphs (a, b, d, and e) each with four infixes and a prefix (c) to (d) indicating the amount of metadata and the relative uncertainty of these metadata sets.

Perhaps the preponderance of faces in the Maya visual language is not accident. Faces can provide the fundamental means of individual visual recognition in visually oriented species. Rather than memorize features of data sets with no spatial arrangements, it may be easier for a user of a glyph-oriented visual language to memorize information encoded as infixes arranged in the form of a face. By remembering the general form of a face as a whole, the user also may be able to remember the data features. Users may find it easier to remember information encoded as facial features than a random or even a sequential ordered array of glyphs and affixes. The face can be conceptualized a pneumonic device. The viability of this approach as a method to promote data comprehension and communication is an open research question.

One possible application of the facial design within a glyph is to arrange a group of four infixes in the shape of a face: nose, mouth and two eyes. The mouth glyph even could be curved up or down to indicate a smile or a frown respectively to indicate something positive or negative in the mouth infix that pertains to the main sign. For example, the infix in the position on the face where the mouth would occur could indicate a data set on friendly or hostile troops with a color code to match. Figure 2 depicts infixes arranged in a configuration that suggests facial features to facilitate information retention and recognition of the information encoded in the face-like arrangement. Figure 2 (b) shows an infix that symbolizes a data set on a neutral entity depicted here as a rectangular "mouth" with vertical hatch marks.

In addition to the fine structure in main signs, affixes can have their own fine structure. Depending on how affixes are defined, they could act as "buttons" that trigger queries or that automatically display the results of pre-retrieved queries stored in the location to which the affix is a pointer. For example, by clicking on any infix encoded as a facial feature, the user could open a data set that pertains to the main sign. The prefix in Figure 2 (c) to the main sign (d) indicates the presence of two metadata sets color-coded to match the oval eye-shaped infixes depicted in the facial glyphs.

One possible application of fine structure in affixes when used as "drill down" to access pedigree metadata [1] is for the affix itself indicate on an approximate logarithmic scale the amount of metadata available. For example, Figure 2 (c) shows a rectangular prefix with two metadata sets color coded to match the infixes in the main sign. Each infix in the main signs (a, b, d and e) of Figure 2 indicates a data set. The prefix (c) indicates that the amount of metadata available to describe the data set symbolized by the eye-shaped infix to the reader's left is ten times the size of the metadata set that describes the data set symbolized by the eye-shaped infix on the reader's right.

Another possible use of an affix is to indicate the uncertainty associated with the metadata. The user could access the metadata sets by clicking on the affix to view pointers to the data sets if the sets are large, or to view the actual metadata in case of a small data set. Figure 2 c displays a possible design for affixes to indicate the relative uncertainty (U 8 vs. U 5) about pedigree metadata in affixes that pertain to the main sign. Color-coding may also be used in glyphs and affixes to convey meaning. For example, data on known hostile forces may be color coded red, as in Figure 2 e, whereas data on forces or a militia suspected of being hostile may be color-coded pink, as depicted in Fig. 2 d.

5.4 Imagery notation

Symbols based on principles and concepts of Maya glyphs could be used to annotate imagery without adding much to the size of the image file. Much information could be encoded in glyphs used for this purpose, thus saving bandwidth. If the image is rich in detail, the margin may be the only place to put the annotation. The reading order of the glyphs that describe the features in the image may be indicated using the Maya method of having a face glyph that points in the direction of the reading order [4].

5.5 Tactical situation assessment language

As mentioned in Section 1, a long-term goal of the TSAT project is to integrate three areas of research: chat, iconic language and wireless gesture-recognition and transmission data gloves. This could lead not only to new and useful results, but it also could open new areas of multidisciplinary research as well. For example, a two-way integration is already in progress with the research into VisChat, which concerns the enhancement of chat with visual icons.

Due to fact that the Mayan scribes did not have access to computers for obvious reasons, the Maya hieroglyphic language itself necessarily remains a visually static language in terms of iconic animation. That not withstanding, we can ask the question, "If the Maya had the means to animate their glyphs, how would they likely use animation to enhance their language?" The study of this question will be worthwhile if it results in useful insights that translate to enhanced applications of modern, emerging, computer-based visual languages.

Research in iconic-language, the data glove and tactical chat come together here with a three-way integration of icons based on the Maya glyphs, and the equivalent representation of certain glyphs in chat and in hand gestures from.

5.5.1 SA report header

An "action glyph" can be found among the Maya glyphs that signify details of the lunar cycle. In this glyph, called "Glyph C," a flat hand holds one of several objects, such as an eye, a head or a skull [5]. The flat hand glyph is designated as "T713" in the modern system of Maya glyph identifiers. This flat-hand glyph is a component of "Glyph C," the eye variant of which is depicted in Figure 3 [5]. In Figure 3, the flat hand is said to be holding an eye, although the part that represents the eye is not entirely obvious.

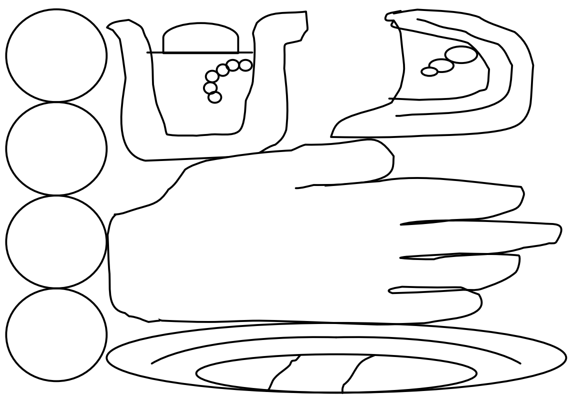


Figure 3. Maya Glyph C with flat hand (T713) and eye [5].

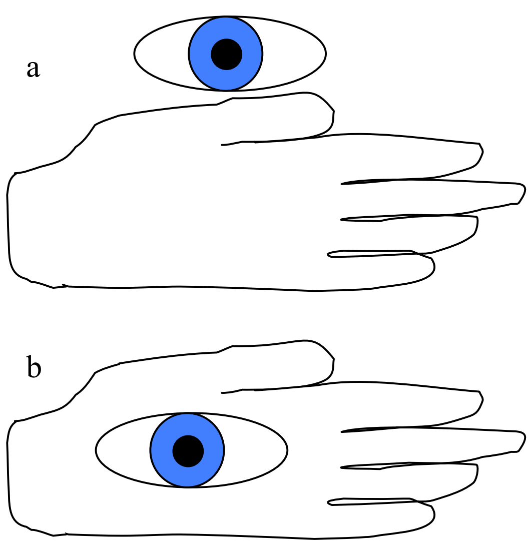


Figure 4. Proposed military flat-hand SA-report header icon candidates based on Glyph C [5] showing eye and flat hand, "Ihand," with the eye as a) superfix and b) as an infix.

In a military chat application, a hand depicted in a manner very similar to Glyph C could be animated and featured in displays where it could reflect and reproduce dynamic gestures from the wireless data glove. These dynamic gestures would interact with the surrounding, possibly static part of an iconic display in the same way in which T713 interacts with the other elements Glyph C.

For example, a very important message in tactical SA is the initiation of a report transmission on SA. The best solution to any problem of this nature will be the simplest to remember, initiate and perform; the least expensive in terms of resources; the most timely, and. the most efficient. An animated dynamic gesture based on the Maya Glyph C with eye [5] can be used for this SA message header. This gesture can be transmitted from a data glove to a computer screen where it can appear either as a dynamic animated icon, a static icon, or in chat, simply as "ihand." It translates into English as "I (**eye**) see what's going on and I have a **hand**le on the situation," or simply "I (**eye**) **hand** you a report." The eye in the hand concept has been used for centuries as a symbol of knowledge and wisdom. It will be easy to remember for this reason.

Figure 4 depicts two simplified versions of Glyph C that could function as SA report headers. The data-glove transmission would proceed as follows. The glove user touches the tip of the thumb to the outer-most joint of the index finger. This forms an eye-shaped aperture between the thumb and finger. This gesture is followed by another gesture in which the hand is made totally flat, similar to the hand depicted in the Maya Glyph C. This series of gestures is repeated three times to signify the start of a situation report.

The advantages of using this report header are many. First, it is simple and easy to remember for obvious reasons. Second, it can be done in any orientation using either the right or left hand. Only one hand is necessary to transmit this "ihand" SA header. Soldiers fighting may have to hide to avoid detecction in awkward positions and orientations where only one hand is accessible for use with the data glove. Moreover, if a soldier has an injured hand or arm, the other side could be used to send the "ihand" SA header gesture. Third, this gesture has the following simple computer representations:

a. Dynamic icons that practically imitate the exact hand gesture,
b. Static icon or glyph showing a flat hand with an eye as a superfix or as an infix,
c. "Ihand" in chat.

5.5.2 Yes and no

Another example is an icon depicting a fist with thumb up inspired by Glyph C with jaguar head or with female head [5]. This icon with glove gesture to match can mean "yes," "positive," or "alive" depending on the context. An icon meaning the opposite can be created based on Glyph C with skull [5]. It can represent "no," "negative," or "dead." The corresponding hand gesture would be a fist with thumb pointing down. The chat counterpart could be "+" and "-" respectively, or simply "yes" and "no."

This protocol also has all the advantages of "ihand" SA header, including the basis for historic usage dating back to the Roman gladiators who lived or died by these gestures.

6. Directions for future research

Additional aspects of the structural principles of Maya hieroglyphic language can be studied. Future research can focus on the identification and test of candidate hand gestures that communicate the specific knowledge glyphs and their affixes. For example, the identification of a set of simple, intuitive, and possibly related hand gestures that can be translated into glyphs and icons used to enhance the COP in C^2 visual displays is an open research question.

7. Conclusion

In summary, modern icon design has not borrowed sufficiently techniques from natural language. Maya hieroglyph script provides a wealth of efficient approaches to glyph design and further analysis of the use of icons to encode information in displays and in imagery.

Acknowledgements

The authors thank the Defense Threat Reduction Agency, the Office of Naval Research and the Space and Naval Warfare Systems Center, San Diego, Science and Technology Initiative for their support of this work. This paper is the work of U.S. Government employees performed in the course of employment and no copyright subsists therein. This paper is approved for public release with an unlimited distribution.

References

[1] M.G. Ceruti, A. Ashfelter, R. Brooks, G. Chen, S. Das, G. Raven, M. Sudit, and E. Wright, "Pedigree Information for Enhanced Situation and Threat Assessment," *Proceedings of the 9th IEEE International Conference on Information Fusion (FUSION 2006),* Florence Italy, July 2006.

[2] S. Fugate, E.W. Medina, L.T. Duffy, D. Magsombol, O. Amezcua, G. Rogers and M.G. Ceruti, "Next-Generation Tactical-Situation-Assessment Technology (TSAT): Iconic Language," *Proceedings of the First Annual Visual and Iconic Language Conference (VAIL 2007)*, University of New Mexico, Aug. 2007.

[3] L. Heinzmann, *Phonetic Picture-Writing,* 2006. http://members.aol.com/leonheinz/english-lautbildschrift/lautbildschrift.htm

[4] Lead Standardization Activity (LSA), Center for Standards (CFS), *Department of Defense Interface Standard Common Warfighting Symbology*, MILSTD2525B, Reston Virginia, Jan. 1999.

[5] J. Montgomery, *How to Read Maya Heiroglyphs*, Hippocrene Books, Inc. New York, 2002.

[6] A.C. van Rossum, *Visual Iconic Language: A Framework for Visual Iconic Languages*, pp. 1-201, Master Thesis, Delft University of Technology, Netherlands, June 2006.

[7] G. Thomas, R. Whitaker and G. Barbier, "Knowledge Glyphs: Iconography to Assist in Command and Control," *Proceedings of the First Annual Visual and Iconic Language Conference (VAIL 2007)*, University of New Mexico, Aug. 2007.

[8] N. Webster, *Webster's New Universal Unabridged Dictionary*, Deluxe Second Edition, (Webster's New Twentieth Century Dictionary), J. L. McKechnie, editor, Simon and Schuster, New York N.Y., 1983.

[9] Wikipedia, *Maya script*, http://en.wikipedia.org/wiki/Maya_script

[10] Wikipedia, *Navajo language*, http://en.wikipedia.org/wiki/Navajo_language

[11] D.R. Wilcox and M.G. Ceruti, "A Structured Service-Centric Approach for the Integration of Command and Control Components," *Proceedings of the IEEE International Conference on Service Computing*, 8-11 July 2008.

Presenting Author Biography

Dr. Marion G. Ceruti is a scientist in the Command and Control Technology and Experimentation Division of the Command and Control Department at SSC-SD. She received the Ph.D. in 1979 from the University of California at Los Angeles. Dr. Ceruti's professional activities include information systems research and analysis for command and control decision-support systems, novel approaches to language for military communications, information systems for chemical and biological defense, sensor-data fusion, and research management. She is the author of over 100 publications on various topics in science and engineering. Dr. Ceruti is a senior member of the IEEE and a member of the IEEE Computer Society, the Association for Computing Machinery, the Armed Forces Communications and Electronics Association (AFCEA), the International Society for Computers and Their Applications, the American Chemical Society, and the New York Academy of Sciences.

A Visual Query Language for Correlation Discovery and Management

Geoffrey M. Draper*
School of Computing
University of Utah

Yarden Livnat†
Scientific Computing and
Imaging Institute
University of Utah

Richard F. Riesenfeld‡
School of Computing
University of Utah

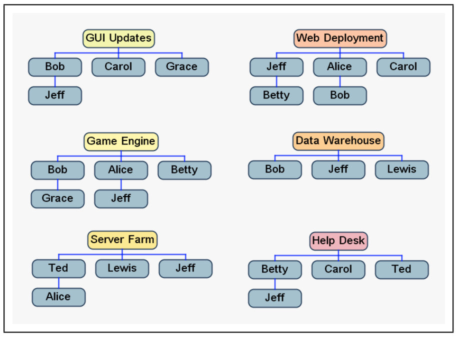


Figure 1: Typical assignment chart.

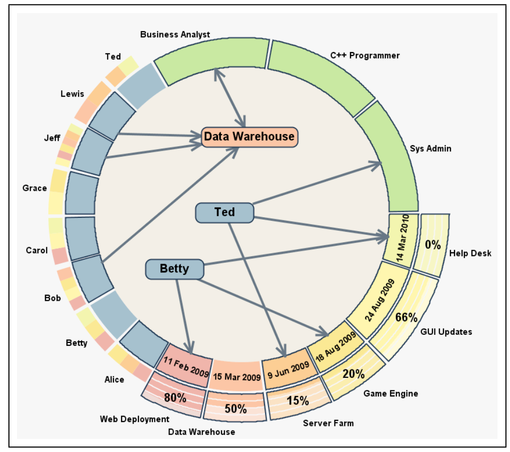


Figure 2: Our proposed visualization.

Abstract

Discerning relationships among entities in a large iconic chart can be a challenging task. As an alternative to traditional chart layout schemes, we introduce an interactive radial query language for simplifying the task of searching for and identifying subtle correlations among data. Our approach allows the user to select which entities to show relationships for, thereby decreasing the cognitive overload associated with static charts. We likewise present a compact visual representation for comparing the differences between two versions of a chart. We have also implemented an intuitive gesture-based interface for creating and removing links between entities, thus enabling users to edit data, not just view it. Our preliminary user trial suggests that users can discover correlations significantly more quickly and accurately with our method as compared to traditional chart representations.

Keywords: Project management, resource allocation, situational awareness, gestural interfaces, visual diffs, social networks.

Index Terms: K.6.1 [Management of Computing and Information Systems]: Project and People Management—Staffing; I.3.6 [Computer Graphics]: Methodology and Techniques—Interaction techniques

1 Introduction

Locating specific information in a large iconic chart can be a challenging task. Traditional charts are both *static* and *accessible* [48], meaning their complete state is visible to the viewer at all times. Consequently, as a chart grows in size, certain tasks become increasingly complex. For one thing, it is difficult to locate specific entities within the chart. Also, it is quite time-consuming to determine relationships among entities in the chart, other than the relationship for which the chart was constructed.

Consider, for example, the "assignment chart" shown in Figure 1. This chart depicts the current projects of a hypothetical small business, along with the names of the employees currently assigned to each project. While useful in some situations, charts such as this are not well suited for answering queries like, *Which projects is a particular individual working on?* Or, *Given two employees, which projects do they work on together?* Such queries are likely to come up whenever a manager needs to evaluate the state of a project or to reallocate resources among projects. Although this information is present on the chart, it is not displayed in an easily accessible format. The situation is further complicated if the manager wants to view a dimension of the data that is not visible on the chart; for example, a list of skill sets required to complete a certain project.

To address these issues, we propose a computer-based visual query language for data sets typically rendered as static charts (Figure 2). Icons representing individual entities are placed around the circumference of a ring, and relationships among multiple entities are shown when an icon is dragged into the ring's interior. In this manner, correlations in the data are displayed on demand, without the need to manually search for them within a chart.

In contrast to traditional chart rendering techniques, our method only shows links for entities selected by the user, helping to reduce visual clutter. However, for each selected entity, every link between that entity and any others are displayed, thus allowing the user to discover relationships that are not readily apparent from static diagrams. Our visual language also supports the ability to add and remove links between entities, as well as visually compare two versions of a data set.

The concepts presented in this paper have applicability to a wide range of problem domains. For the purposes of our initial prototype, however, we have chosen to focus on the specific problem of helping project managers allocate their resources effectively. We discuss other possible application domains in Section 5.

Although project managers have many variables to coordinate, we have chosen to focus on three: the human *resources* on his or her team, the *projects* or tasks assigned to the team, and the *skill*

*e-mail: draperg@cs.utah.edu
†e-mail: yarden@sci.utah.edu
‡e-mail: rfr@cs.utah.edu

sets or proficiencies possessed by the team members, and required by the projects.

The remainder of the paper is organized as follows. We first summarize related work in the areas of project management visualization, visual query languages, and radial visualization. We then present a detailed overview of our visual query language, including its syntax for both standard queries and difference queries. The next section mentions some other application domains to which our visualization could be adapted, suggesting the generality of our technique. We next describe our prototype software implementation, with an emphasis on its editing features and level of detail controls. Finally, we report on the findings of our informal user study, and conclude by proposing directions for future work.

2 Related Work

Our work arises from the needs of a number of fields, namely, project management, visual query systems, and radial visualization. We briefly review some of the related work in each of these areas.

2.1 Project management

During the late 1980s and extending into the 1990s, there was a surge of interest in software-based tools for project management. This led to the development of a number of systems [9, 36, 40, 44], with Microsoft Project eventually becoming a *de facto* standard [37]. As Microsoft Project aims to be a general purpose system, other project management tools also exist as specialized solutions to niche problems [19, 25, 39].

There has also been a great deal of research incorporating aspects of information visualization into project management. The most well-known visualizations for project management are the PERT chart and the Gantt chart. One of the earliest examples of computer-generated visualizations for project management is the *progress diagram* [33] which displays some of the same information as a Gantt chart, but resembles a tree structure in appearance.

Szekely et al. [57] use the established visualization technique of *star fields* [1] to assist in planning military missions. Wang and Lillehagen [59] employ a variety of known diagram types (*swim lanes*, *trees*, and others) in their system, but do not introduce any new visualizations per se. Stamey and Honeycutt [53] apply a *quad chart* to the domain of project management. Kontio et al. [30] use a form of directed acyclic graph (DAG) to visualize risks associated with projects. The *SSD graph* [32] is an attempt to visualize both the structure and the state of a project.

Halin and Hanser [18] present a novel visualization system to show relationships between resources and tasks. Of all the systems reviewed herein, this one perhaps is most closely related to our work. However, we take a somewhat different approach, and focus, in particular, on interactivity and ease of use.

2.2 Visual query systems

Making data analysis easier has been a subject of research for several decades. Cammarano et al. [3] make the observation that most user interfaces for data analysis take either one of two approaches. In the first case, the user interface simplifies the formulation of the query; in the second case, the user interface helps to visualize the results.

Catarci et al. [5] provide a comprehensive, if slightly dated, survey of the literature in the first category. Of particular interest is the problem of querying semistructured or unstructured data, such as websites or emails. Sinha and Karger [51] propose a system for aiding in navigation of semistructured data sets by suggesting navigation hints to the user. Trigoni [58] lets the user refine a query over time by gradually disclosing the underlying data one part at a time. Goldman and Widom [17] propose a method for exploiting similarities among pages in the same website to perform more effective queries. Polyviou et al. [46] describe an interface for performing database queries based on the ubiquitous filesystem browser interface. The VisTrails system [49] makes use of provenance data to maintain a history of past queries for creating visualizations.

In the second category, there are many systems which offer a direct-manipulation interface for browsing the results of a query. Furnas and Rauch [15] as well as Stonebraker [56] present canvas-based visualizations that support zooming and panning. XmdvTool [60] facilitates the creation of standard statistical graphs to display query results. Visage [47] by Roth et al. is a highly interactive direct manipulation system that uses a variety of graphing techniques to communicate results to the user. In addition, Keim [27] and de Oliveira and Levkowitz [8] present surveys of many database visualization techniques.

In contrast, the Polaris system [55] incorporates both a novel query interface mechanism and an integrated visualization. Based on the well-known metaphor of a pivot table in a spreadsheet, Polaris allows users to view correlations in the data with respect to any attribute in the data set. Our work is similar in that we also provide a streamlined interface for performing queries and for *in situ* visualization of the query results.

2.3 Radial visualization

Our design metaphor is based on the radial layout, a charting technique whose tradition spans back hundreds of years, and is an increasingly popular metaphor in contemporary information visualization science. Pre-digital examples of radial information layouts include William Playfair's 1801 invention of the pie chart [45, 52], and Florence Nightingale's rose diagrams [41] for communicating sanitary conditions in British military hospitals during the Crimean War. In the mid-twentieth century, Northway used radial diagrams to track the social behaviors of gradeschool children [43].

Much of the recent work in radial user interfaces traces its lineage to research in graph layout algorithms for computer graphics [21]. These algorithms, in turn, have inspired techniques for visualizing multivariate data. Many such designs involve positioning data points as nodes on the spokes on a wheel [20, 22]. In these visualizations, the center point of the canvas holds some semantic meaning, and the distance of each node from the center shows a relationship relative to it. A recent example of this is the DataRose from [11].

In contrast, a second variety of radial visualization (called *radial space filling* or RSF [54]), the data points are typically arranged in compact concentric rings [29] or in a spiral [2, 4], and rendered so as to form a circle. Each ring contains a different category of data. Examples include polar treemaps [24], fan charts [10], and Radial Traffic Analyzer [28].

Another general category of radial visualization arranges the data points around the circumference of a ring, while reserving the interior of the ring for other data. Correlations among data points are often rendered as lines between nodes on the circumference and nodes in the interior. Examples include *Daisy* [7], *NetMap* [16], and *VisAlert* [34].

3 Visualization Concepts

In our visualization metaphor, all available entities (projects, resources, and their related skill sets) are arranged in a ring. This egalitarian approach places all variables on equal footing; none receives greater emphasis than the others until it is selected by the user. From a user interface perspective, this visualization has the advantage that each entity is located within close proximity to its neighbors [13], rather than being distributed over a large area as in traditional charts.

3.1 Interactive relationship discovery

The mode of interaction is perhaps best communicated by example. We present a scenario of how a project manager

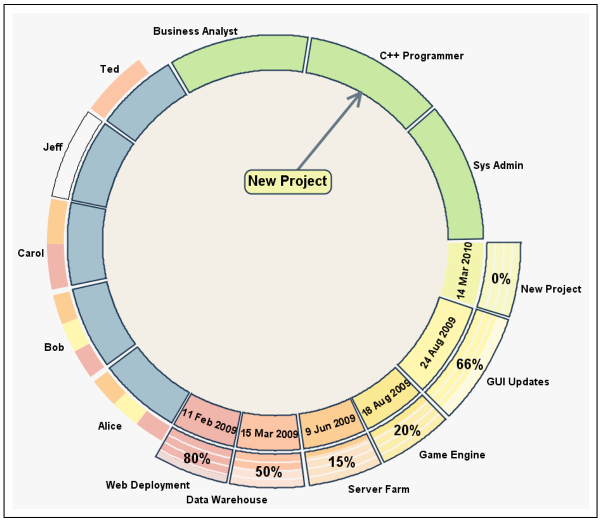


Figure 3: Relations between a project and a required skill.

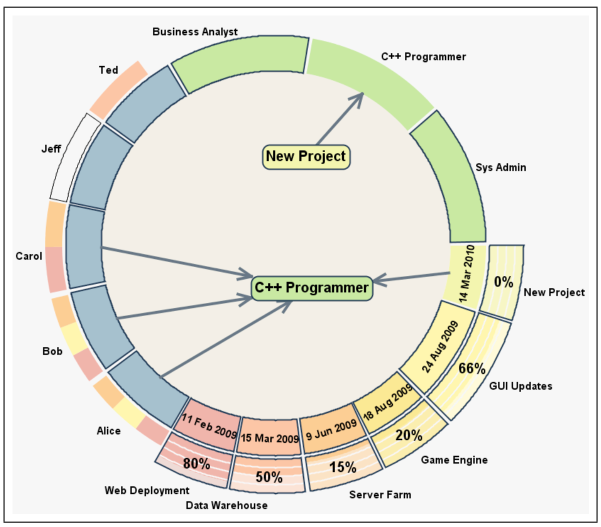


Figure 4: Relations between a skill, the projects requiring it, and the resources possessing it.

might use our visual query language to handle a typical task. For a more visual example, we refer the reader to a demo video (in QuickTime format) which may be downloaded here: `http://www.cs.utah.edu/~draperg/research/` .

Suppose that Jane, a team manager, receives a request for her team to take on a new project. The procedure described below outlines the steps that Jane might take in choosing which resources to allocate to this project, using our visualization.

Step 1 Jane selects the proposed project and brings it into the interior of the ring to discern its staffing requirements. Jane sees that this project requires a C++ programmer (Figure 3).

Step 2 Next, by introducing the skill labeled "C++ Programmer" into the ring's interior, Jane recalls that there are three people on her team who have this skill: Alice, Bob, and Carol (Figure 4).

Step 3 Jane decides to look further into Bob's profile. Placing Bob into the ring, she sees which projects he is currently involved

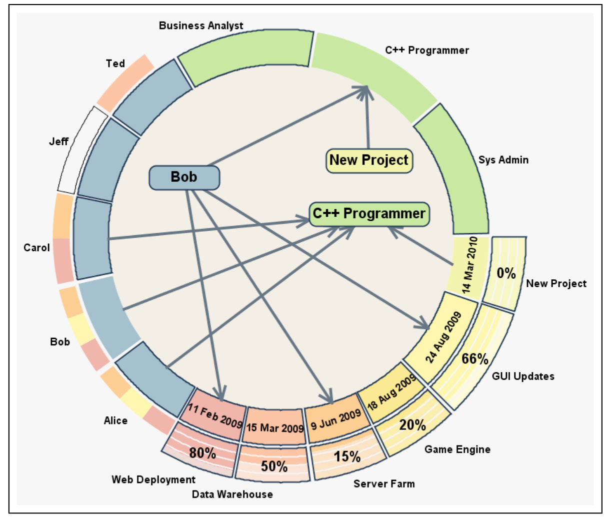


Figure 5: Relations between a resource, the skills it possesses, and its assigned projects.

with: Web Deployment, Server Farm, and GUI Updates (Figure 5).

By examining the current commitments of Bob and others, Jane can make an informed decision regarding whether to accept this new project, and how best to distribute the workload among her team members.

In addition to the basic interaction metaphor outlined above, our visualization's core strength is in allowing the user to uncover complex relationships within the data, beyond what can be effectively displayed in traditional charts. When multiple icons are placed in the interior of the ring, all known relations between these entities and the other entities are made visible. Whenever two or more arrows converge on the same point on the circumference (Figures 2 and 14), that is a clear visual indicator that these entities have a relationship in common.

For example, if two people consistently work on the same set of projects together, that might signal to their manager a need for them to collaborate more with others. Likewise, when multiple skills are placed inside the ring at the same time, the user can immediately see the distribution of skills among the team, and whether there are enough resources to fill the requirements of the team's assigned projects. Or, if multiple projects are selected, the relationships of skills and people to projects become visible. In each of these examples, there is admittedly some overlap on the information being presented. The key is the flexibility of our approach — the user is free to select the method that best lends itself to the question he or she is trying to answer.

3.2 Space-saving visualization

One advantage of our method is its conservative use of space. In a standard assignment chart, every time an entity participates in a relationship, a separate instance of that entity must be included in the chart — thus linearly increasing its overall size. In our visualization, the size of the ring is constant, and each entity is shown at most twice: once along the circumference, and once inside the ring *if* that entity is currently selected by the user.

For example, consider the chart shown earlier in Figure 1. For each project to which Betty is assigned, a separate icon for Betty must be added to the chart. Contrast this with the visualization in Figure 2: Betty's multiple assignments are still clearly indicated; however, Betty's name appears only once in the ring's interior. Perhaps more importantly, the overall "footprint" of the visualization

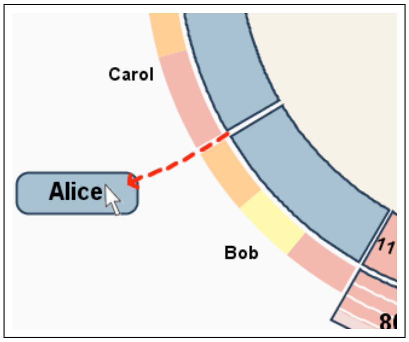


Figure 6: Control level of detail by removing unwanted entities from the ring.

does not change, regardless of how many relationships are shown.

3.3 Scalability and level of detail

One known weakness of our current approach is that the visualization does not scale well as the number of nodes exceeds about 50 or 60. This is due to the constraint of fitting all nodes within the limited space around the ring's circumference. This is an issue shared by most radial visualization systems in general. Nonetheless, we offer two suggestions for ameliorating the scalability problem in our visual language.

First, any icons that are not of current interest to the user may be dragged outside of the ring (Figure 6). This not only reduces crowding around the circumference of the ring, but reduces the number of relations that are displayed for icons inside the ring, because no lines are drawn for nodes that are not currently on the circumference. Whenever the user wishes to include this entity again, its icon can be dragged back onto the circumference of the ring. In this manner, the amount of detail displayed about any resource, project, or skill is configurable on the fly — allowing the user to view as much or as little information about a given entity as desired.

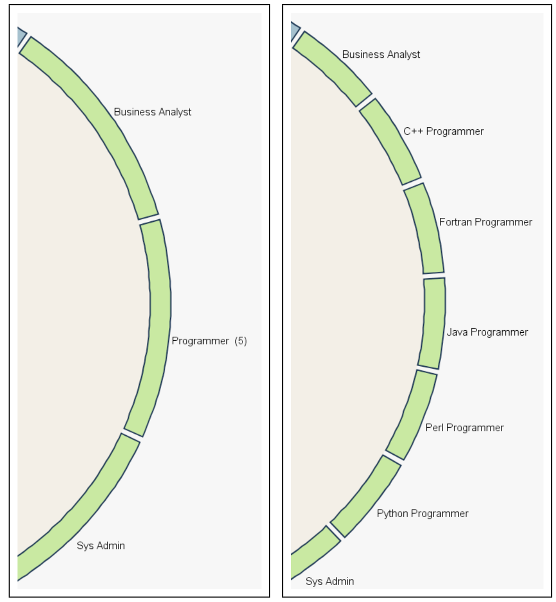
A second way to address the visualization's scalability is through aggregating similar entities. For example, if a team contains people with skills in several different programming languages, all these skills could be aggregated into one composite skill, labeled "programming" (Figure 7(a)). At the user's option, this skill could be expanded to reveal the individual skills (Figure 7(b)). Similarly, individual resources could be aggregated into teams of resources, and projects could be grouped according to categories of related projects. Furthermore, these aggregate entities may be combined into still more generalized entities until a desired level of abstraction has been achieved. In this manner, scalability is not so much limited by the size of the circle, but by the level of detail the user wishes to see [42].

In practical usage, a manager over a team of engineers would likely want to see his or her resources, projects, and skills at the lowest level of granularity. A middle-tier manager or executive, on the other hand, might prefer a broader, less detailed overview of what the organization as a whole is doing.

Our prototype system, described in Section 6, currently supports the first of these two proposed techniques. The second is an area for future work.

4 Difference Visualization

Once one succeeds at visualizing a single data set, a logical next step is to visualize the differences between two data sets. By far, the most common domain for difference-finding is in text documents. A *de facto* tool for file comparison is `diff`, which has been included in the Unix operating system since 1974.



(a) A composite skill. (b) Detailed view of skills.

Figure 7: Composite and expanded entities.

However, `diff`'s output, which uses '>' and '<' symbols to signal differences between files, is not obvious for beginning users. Current versions of `diff` optionally support more intuitive symbols, such as '+' and '–', to denote insertions and deletions within the text.

Although the algorithmic problem of isolating differences between two text files is essentially solved, comparatively little work has been done on the problem of finding differences in other kinds of data, e.g. graphs and charts, although there are important exceptions [12, 14, 31, 49].

We propose two complementary visualizations for comparing different versions of a chart. The first is based on a traditional chart-based representation, while the second uses an interactive radial approach, similar to the one presented above.

4.1 Chart Visualization

Our visual diff metaphor is based on the observation that between any two versions of a chart, a link between two nodes can either be:

- present in both charts,
- present in the first chart and absent in the second chart, *or*
- absent in the first chart and present in the second chart.

Our visualization is inspired by modern implementations of `diff` that use the '+' and '–' characters to signal changes in the data. A link that is present in the first chart, but absent in the second chart is shown by drawing a '–' in the icon adjacent to the link. A link that is absent in the first chart, but present in the second chart, is similarly depicted by a corresponding '+'. Figure 8 shows an example of our representation. This visualization is significantly more compact than simply putting the two versions of the chart side by side.

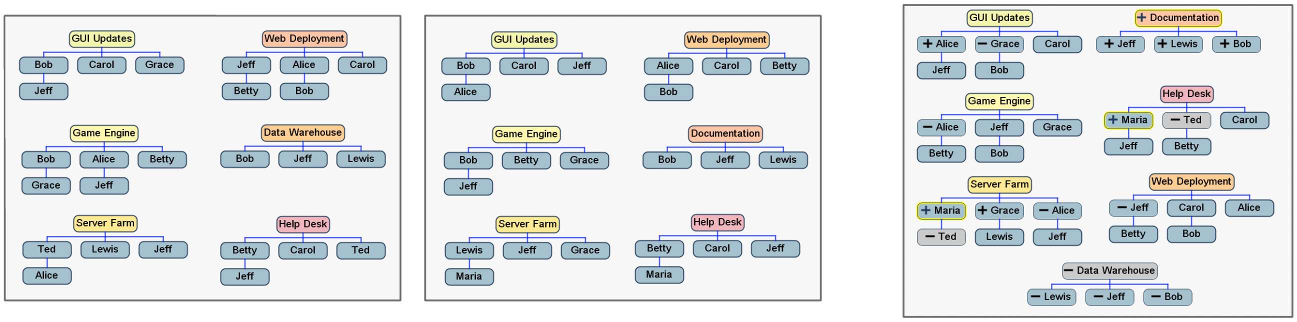


Figure 8: The leftmost chart represents Version 1 of a chart; the middle chart is Version 2. The chart on the right is a “Visual Diff” of the two versions.

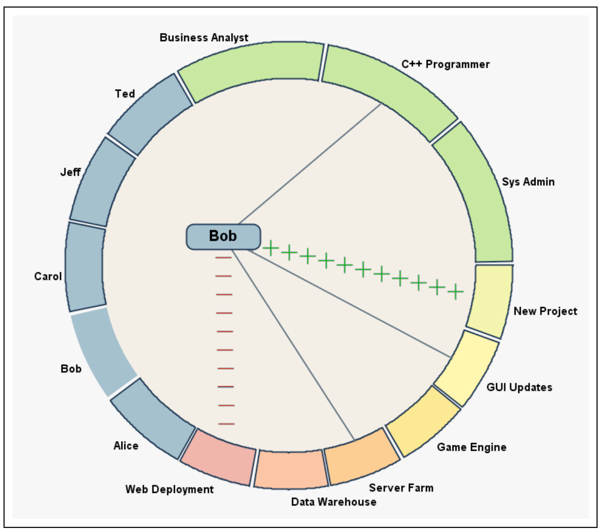


Figure 9: Interactive radial representation of a “Visual Diff” chart.

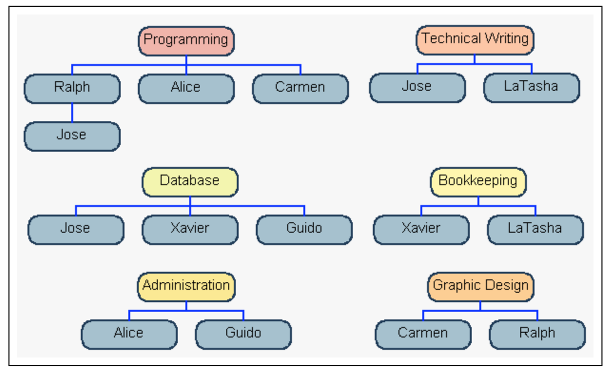


Figure 10: Graph showing skill distribution among team members

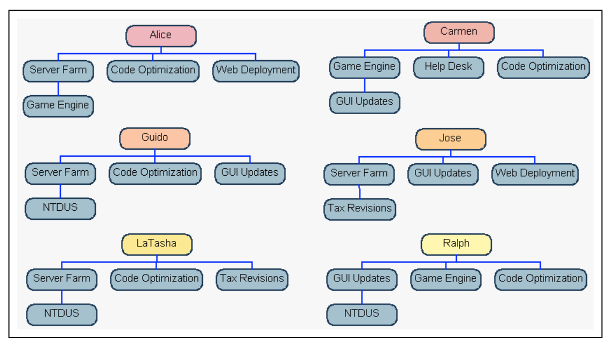


Figure 11: “Inverted” assignment chart

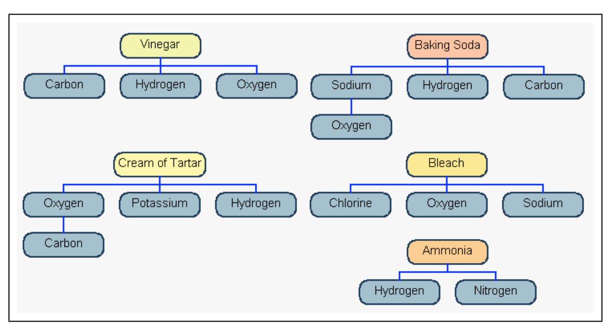


Figure 12: Graph of ingredients in household chemicals

4.2 Radial Visualization

Our second representation for finding differences is an interactive visual language based on the radial query metaphor introduced in Section 3. As before, icons representing the resources, tasks, and skills are arranged in a ring. Icons shown in yellow highlight are those that are new to the second version of the chart. Icons shown in a muted gray color are those that appeared in the first version of the chart, but were absent from the second. When the user drags the node’s icon into the ring’s interior, links are rendered as lines between the interior node and one or more nodes on the circumference of the ring. Now, however, the line style indicates the status of a link. An old link, that is, a link that existed in the first chart but not the second chart, is rendered as a sequence of ‘–’ symbols. A new link, one that is unique to the second chart, is rendered as a line of ‘+’ signs. Predictably, a link that exists in both versions of a chart is rendered as a solid line (Figure 9).

5 “One-to-Many” Charts

Although we have focused thus far on the application domain of assignment charts, project management is not the only setting in which charts like this may be used. An assignment chart is simply a specific instance of a more general class of charts which we call “one-to-many” charts. In an assignment chart, for example, each (one) project is linked to all the (many) people who are assigned to it.

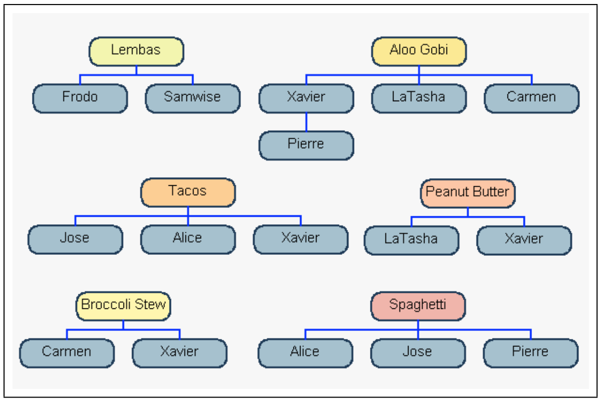


Figure 13: Graph of individuals' culinary preferences

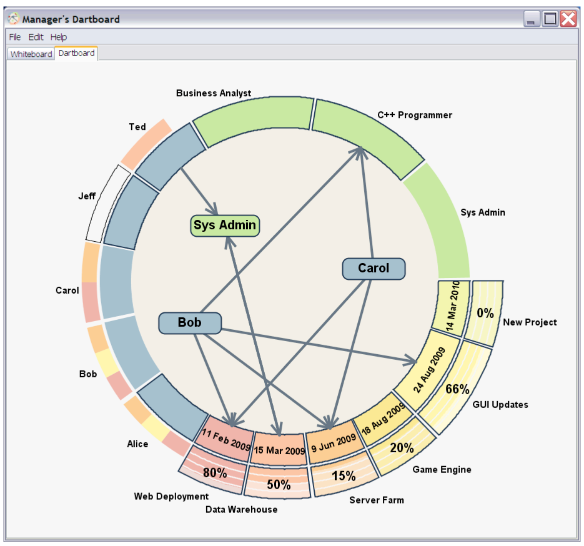


Figure 14: Screenshot of the prototype system, *Manager's Dartboard*.

Although assignment charts are a common type of one-to-many chart, they are certainly not the only kind. Consider a "skill chart", showing which skills are possessed by individuals within an organization (Figure 10). Or an "inverted assignment chart" might show the tasks to which a resource is assigned, as opposed to which resources are assigned to a task (Figure 11). Stepping away from our "project management" motif, one could imagine a chart that displays the names of atomic elements alongside which household chemicals they are found in (Figure 12), or a chart that shows which people enjoy certain types of food (Figure 13).

As the above examples show, this class of chart has a variety of applications. The interactive visual language presented in this paper could be easily adapted to any one of these domains.

6 Manager's Dartboard

As a proof of concept of the visualization methodology presented herein, we have implemented a prototype software system called *Manager's Dartboard* (Figure 14).

In order to free the user from the burden of frequently referring to another part of the interface (or to another program entirely) for supplementary information, we have tried to make the icons as informative as possible without making our interface overly busy. One purpose for doing this is to allow the user to maintain focus

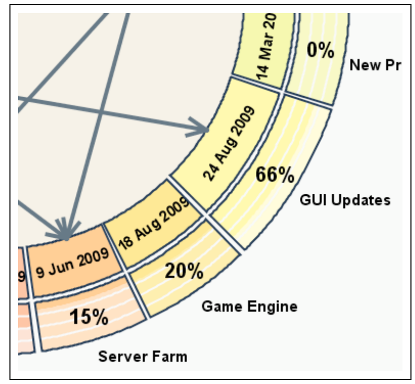


Figure 15: Project icons.

on the visualization task, rather than having to perform a time-consuming (and potentially distracting) "context switch" to find the information elsewhere.

As such, Manager's Dartboard supports multiple levels of inquiry:

1. The icons themselves provide basic information about a particular entity at first glance.

2. More detailed information is available via a tooltip if the user hovers the cursor over the icon.

3. The menu system provides access to dialog boxes allowing the user to view and edit all of an entity's attributes.

We feel that this *details on demand* approach [50] strikes a fair balance between the opposing ideals of maintaining a clean user interface, and providing detailed information on specific entities as required by the user.

6.1 Project Visualization

In our system, each project possesses a unique 4-tuple of attributes: *name*, *end date*, *percent complete*, and *priority*. We designed the project icon so that each of these attributes is immediately visible without further inspection.

First and most simply, the *name* of the project is printed next to the icon. Also, the expected *end date* for each project is displayed prominently in the icon's interior. Furthermore, the icons are positioned around the ring counterclockwise, sorted in order of the project's end date. In addition, the icon also shows the progress of the project in terms of *percent complete*. The percentage is shown textually in bold type, and graphically as a curved progress bar. Finally, the *priority* of a project is suggested by the icon's radial magnitude (i.e. its "length") relative to the other project icons (Figure 15), with higher priority projects occupying more space than the others.

In this manner, the user can quickly scan the list of projects to see which projects are due first, their relative priority, and how far they are from completion. This helps the project manager readily discern which projects may be in need of further attention.

6.2 Resource Visualization

The resource icons are designed to let the user quickly see both how many projects a resource is assigned to, and — perhaps more significantly — what proportion of the resource's time is allocated to each project. A resource's assigned projects are represented as a band of colored sectors along the radial exterior of the resource icon. We call this band the *project previewer* because it provides the user a hint of how busy each resource is, without needing to drag

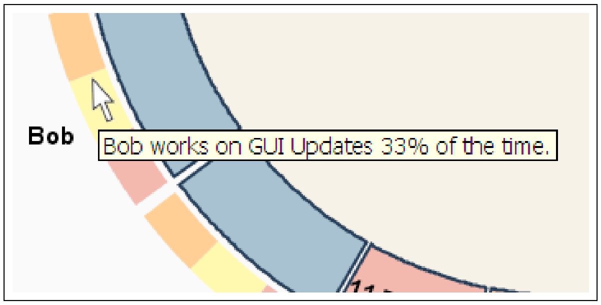


Figure 16: Hovering the cursor over an icon reveals additional details.

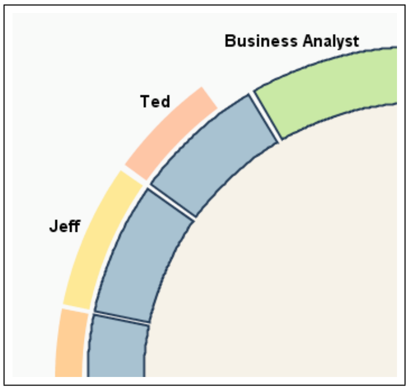
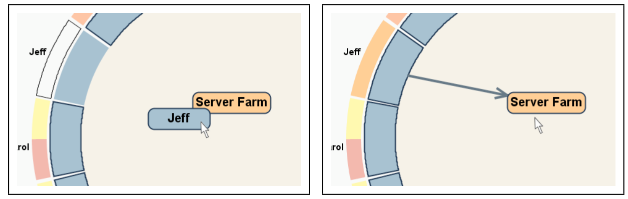


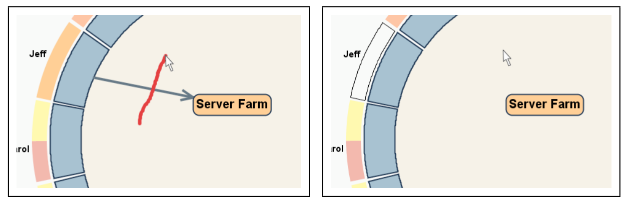
Figure 17: The size of the “project previewer” shows the comparative capacity of resources.



(a) Drag an icon atop another icon to associate them.

(b) The relation is created.

Figure 18: Interface for creating relations.



(a) The “cutting” gesture.

(b) The relation is broken.

Figure 19: Interface for removing relations.

the resource icon into the ring’s interior. The color of each sector in the project previewer corresponds to the color of the corresponding project icon elsewhere on the ring’s circumference. Furthermore, the relative size of a sector as compared to its neighbors is a visual indicator of how much of the resource’s time is devoted to the given project. To reduce the amount of visual clutter in the display, we do not print the names of the projects alongside the project previewer; however, this information is available via a tooltip when the user hovers the cursor over the project previewer (Figure 16).

Even on a highly talented team, no two resources are likely to have the same levels of experience or competency, even though there may be overlap in their skill sets [26, 35]. For this reason, our system allows resources to be assigned a “capacity” percentage to reflect their level of ability relative to other resources. Visually, this is reflected by the length of the project previewer relative to the resource’s icon. For example, if Ted is a newer employee, his capacity might be 75% of Jeff’s, who has more experience. Thus, his project previewer only takes up three-fourths the length of his resource icon, whereas Jeff’s runs the full measure (Figure 17). Note that this does not prevent Ted from being assigned as many projects as Jeff or anyone else; however, the smaller size of his project previewer serves as a visual reminder to managers that he has proportionally less capacity to offer to his projects than do perhaps his counterparts.

6.3 In-place editing

In keeping with our goal of giving the user maximum flexibility, we have also implemented a series of gestural interfaces to allow the user to edit relationships among entities directly within the display. The user may assign resources to projects, indicate skill requirements for projects, associate resources with their respective skills, as well as unassign or dissociate any of the above relationships. The gesture for associating two entities is accomplished by the familiar *drag and drop* metaphor.

This gesture is illustrated in Figure 18. Figure 18(a) shows the user dragging the icon for a resource, Jeff, atop the icon for project Server Farm. Alternatively, the icon representing the project could have been dragged atop the icon representing the resource. When the link is created, the dragged icon disappears and an arrow is rendered between the icon inside the circle and the icon on the circumference, as shown in Figure 18(b). The color of Jeff’s project previewer likewise changes to show that Jeff, who was not assigned to any projects, is now associated full-time with the Server Farm project.

The gestures for associating projects and resources with skills are analogous — simply drag and drop the icons representing the desired entities atop each other, and the system creates the underlying association.

6.3.1 Sketch-based Interface

Motivated by the interface of [23], the gesture for removing a relation between two entities consists of drawing a straight line which intersects the arrow connecting two entities. For example, in Figure 19(a), the user draws a line bisecting the arrow connecting the icons for resource Jeff and project Server Farm. In Figure 19(b), notice how the color of the resource’s project previewer changes to reflect the fact that Jeff is no longer assigned to any projects.

To reduce the chance of error, our system requires that the input gesture for removing an association between entities be an approximately straight line rather than a curved one. We do this by comparing the total length of the *ink stroke* with the distance between its endpoints. A gesture is considered straight if and only if it satisfies the conditions in Equation 1 below.

Let $\rho_{1..n}$ be the vertices of the gesture as inputted by the user.

Let $\delta(i, j)$ be the Euclidean distance between vertices ρ_i and ρ_j.

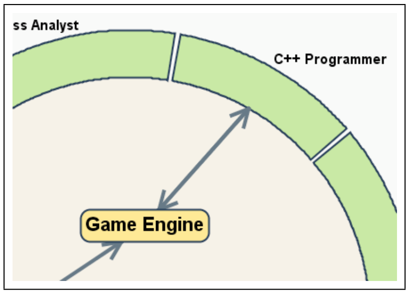


Figure 20: A project whose assigned resources match its skill requirements.

Let τ be the threshold above which the line is considered straight.

$$\frac{\delta(1,n)}{\sum_{i=1}^{n-1}\delta(i,i+1)} \geq \tau \qquad (1)$$

In our experience, setting $\tau = 0.8$ is sufficient for determining the straightness of a line.

6.4 Directed Relations

Our system renders connections between icons as arrows instead of undirected line segments. This was a conscious design choice, intended to make explicit the nature of the relationship between entities. Specifically, arrows are drawn from resources to projects to show that resources work on projects, rather than the inverse. We render arrows from projects to skills to show a dependency relationship: projects require skills. Arrows are drawn from resources to skills to show that resources possess skills. These design decisions are somewhat subjective; other valid mappings could likewise be employed. However, for consistency of presentation, we have adhered to these guidelines for our prototype implementation.

For convenience, we draw a double-headed arrow between a project and a skill whenever a resource possessing a certain skill is allocated to a project requiring that skill (Figure 20).

In our system, we have opted to draw arrows only between an icon on the circumference of the ring and another icon inside the interior of the ring. We do not draw arrows between two icons that both lie on the circumference of the ring. Contrast this to the approach used by [7] and [16]. Similarly, we do not draw arrows from one icon in the interior of the ring to another icon also in the interior of the ring. These decisions were made in order to reduce the amount of "noise" in our visualization, and in light of the fact that no additional information is conveyed by displaying an individual relation more than once.

6.5 Whiteboard

In addition to the "Dartboard" interface which demonstrates the concepts of our research, our software prototype also includes an alternate "Whiteboard" user interface, based on the familiar notion of an assignment chart. The user can switch between the two views by selecting their respective tabs. Both the Whiteboard and the Dartboard views operate on the same data, albeit via a different user interface. When a change to the data is made in either view, that change is automatically reflected in the other.

Whiteboard is little more than a computerized version of an assignment chart that one might find on a typical manager's whiteboard (Figure 21). It allows the user to associate projects and resources by dragging and dropping icons onto a blank panel. Placing an icon near another icon associates the two; similar behavior can be found in commercial diagram-drawing software like Microsoft Visio [38].

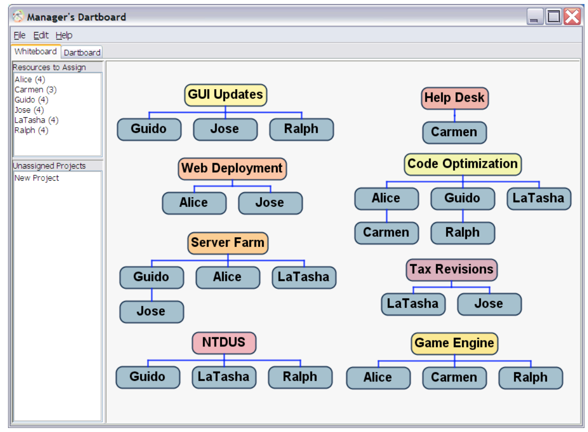


Figure 21: Alternate "Whiteboard" interface

Its convenience of use notwithstanding, Whiteboard was never intended to be especially innovative. Indeed, we implemented Whiteboard solely as a benchmark against which to compare Dartboard, which is the actual focus of our work. For simplicity, the Whiteboard is missing some features that are supported by the Dartboard interface; most noticeably, the ability to manipulate skill attributes.

6.6 Implementation Notes

Our prototype is written in Java. We use the Swing API for the traditional user interface elements (such as the menu system and dialog boxes) and Java2D for the graphics.

Having successfully tested Manager's Dartboard on Linux, Solaris, and Windows platforms, we believe it should run equally well on any platform for which a Java 1.6-compliant virtual machine exists.

7 User Evaluation

To gauge the effectiveness of our visual language, we conducted an informal preliminary user trial consisting of 22 participants, primarily graduate students in computer science. Our goal was to compare the speed and accuracy with which users are able to identify relationship information from both a traditional assignment chart and our visualization. Our participants all performed the test using the same computer, a 2.5GHz Pentium4 PC with 500MB of memory, running Red Hat Linux 9. The screen resolution was set to 1600x1150 pixels.

The trial was in two parts. During Part 1, we showed an assignment chart to the participants, and asked them seven questions based on the chart. The chart was on-screen during the question-answering period so that the participants could refer to it as needed. For Part 2, we introduced them to the Manager's Dartboard system and gave them time to familiarize themselves with its interface. We then presented the participants seven additional questions similar to those in Part 1, letting them use our software to explore the data. Manager's Dartboard displayed all the same resources, projects, and skills as in the assignment chart from Part 1, although we did scramble the project and skill assignments to prevent participants from parroting back the same answers as in Part 1.

The questions consisted of a combination of simple and compound queries, namely:

1. Which projects is Cindy working on?
2. Which projects are both Jeff and Daniel working on?

3. Who is assigned to the greatest number of projects?

4. Which people are assigned to work on exactly three projects?

5. Which one person is assigned to all of these projects: GUI Updates, Web Page Update, Workshop?

6. Which people always work together on the same projects?

7. What skills does Alice have?

Although we initially hypothesized that our visualization's chief advantage would be speed, not accuracy, we were pleasantly surprised to find that the participants not only answered the questions 13% faster on average using Manager's Dartboard, but that they also made an average of 80% *fewer mistakes*. We interpret this as a sign that our method for visual correlation offers a strategic advantage over traditional charts — especially when applied to mission-critical application domains where misinterpretations of data may be costly.

Subjective feedback from users was likewise positive. Upon completing the questions, the majority of participants commented on how much easier Manager's Dartboard was to use than a static chart, especially for determining working relationships among the resources, projects, and skills. While the informal nature of our study makes it difficult to draw definitive conclusions, we are nonetheless encouraged by these preliminary results, and we would like to validate them with additional user studies in the future.

8 Further Work

In our prototype's current state of development, all projects have deadlines. However, some tasks, termed *workflows* by Craven and Mahling [6], are ongoing, repetitive assignments and do not have specified end dates. Examples of workflows could be maintenance, technical support, or committee assignments. We would like to implement support for workflows in the future.

We also plan to expand the ways in which entities can be related to one another. Currently, only entities of different types may be associated together: projects with resources, resources with skills, etc. Additional semantic information could be conveyed by associating entities of the same type. For example, affiliating one resource to another could perhaps express a mentor/trainee relationship, whereas linking one project to another project may designate a chronological or dependency relationship.

Furthermore, we intend to incorporate additional gesture-based commands into the user interface. For example, the proportion of a resource's time that is spent on a given project is currently editable via the menu system and dialog boxes. In the ideal situation, this attribute could be edited by direct manipulation within the visualization canvas.

9 Conclusion

In this paper, we have presented a novel method for visualizing and editing resource allocation data in the domain of project management. Our current prototype supports three types of entities; namely projects, resources, and skills; however, our visualization concept is general enough to accommodate many more.

The main contributions of this paper are:

- a highly interactive canvas for querying multivariate data,

- a gesture-based interface for adding and removing links between nodes, and

- two complementary visual schemes for comparing the differences between two versions of a data set.

Our visual language aims to be a general-purpose strategy for discovering relationships among disparate entities in a wide variety of data sets. Users report that this interface is satisfying to use; furthermore we have found that complex relationships within a set of data can be identified more quickly and accurately using our method than with standard charting techniques.

Acknowledgements

The authors wish to thank Glenn McMinn, CEO of Delcam USA, and Tom McCollough, VP of software development at Delcam USA, for their helpful feedback on an early prototype of this system. We also thank all the participants in our preliminary user study.

This work was supported in part by the Scientific Computing and Imaging Institute at the University of Utah, and by the Air Force Research Laboratory in Rome, NY, under IC-ARDA contract no. F30602-03-C-0257.

References

[1] C. Ahlberg and B. Shneiderman. Visual information seeking: tight coupling of dynamic query filters with starfield displays. In *CHI '94*, pages 313–317, 1994.

[2] E. Bertini, P. Hertzog, and D. Lalanne. SpiralView: Towards Security Policies Assessment through Visual Correlation of Network Resources with Evolution of Alarms. In *IEEE Symposium on Visual Analytics Science and Technology (VAST 2007)*, volume 2, pages 139–146, Oct. 2007.

[3] M. Cammarano, X. L. Dong, B. Chan, J. Klingner, J. Talbot, A. Halevy, and P. Hanrahan. Visualization of heterogeneous data. In *InfoVis 2007*, pages 1200–1207, 2007.

[4] J. V. Carlis and J. A. Konstan. Interactive visualization of serial periodic data. In *UIST '98: Proceedings of the 11th annual ACM symposium on User interface software and technology*, pages 29–38, New York, NY, USA, 1998. ACM Press.

[5] T. Catarci, M. F. Costabile, S. Levialdi, and C. Batini. Visual query systems for databases: a survey. *Journal of Visual Languages and Computing*, 8(2):215–260, 1997.

[6] N. Craven and D. Mahling. Goals and processes: a task basis for projects and workflows. In *COCS '95: Proceedings of conference on Organizational computing systems*, pages 237–248, New York, NY, USA, 1995. ACM Press.

[7] Daisy Analysis Ltd. Daisy 2003. http://www.daisy.co.uk/. Accessed 14 November 2007.

[8] M. C. F. de Oliveira and H. Levkowitz. From visual data exploration to visual data mining: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 9(3):378–394, July 2003.

[9] D. Dixon. Integrated support for project management. In *Proceedings of the 10th international conference on Software engineering*, pages 49–58, 1988.

[10] G. M. Draper and R. F. Riesenfeld. Interactive fan charts: A space-saving technique for genealogical graph exploration. In *8th Annual Workshop on Technology for Family History and Genealogical Research*, Provo, Utah, USA, 2008. Brigham Young University.

[11] N. Elmqvist, J. Stasko, and P. Tsigas. DataMeadow: A Visual Canvas for Analysis of Large-Scale Multivariate Data. In *IEEE Symposium on Visual Analytics Science and Technology (VAST 2007)*, volume 2, pages 187–194, Oct. 2007.

[12] C. Erten, P. J. Harding, S. G. Kobourov, K. Wampler, and G. Yee. Graphael: Graph animations with evolving layouts. In *International Symposium on Graph Drawing*, 2003.

[13] P. M. Fitts. The information capacity of the human motor system in controlling the amplitude of movement. In *Journal of Experimental Psychology*, pages 381–391, 1954.

[14] D. Forrester, S. G. Kobourov, A. Navabi, K. Wampler, and G. Yee. graphael: A system for generalized force-directed layouts. Available at http://graphael.cs.arizona.edu/graphael/, 2005. Accessed 22 August 2007.

[15] G. W. Furnas and S. J. Rauch. Considerations for information environments and the NaviQue workspace. In *INEX Workshop*, pages 79–88, 2004.

[16] J. Galloway and S. J. Simoff. Network data mining: methods and techniques for discovering deep linkage between attributes. In *APCCM '06: Proceedings of the 3rd Asia-Pacific conference on Conceptual modelling*, pages 21–32, Darlinghurst, Australia, Australia, 2006. Australian Computer Society, Inc.

[17] R. Goldman and J.Widom. Interactive query and search in semistructured databases. In *International Workshop on the Web and Databases, WebDB'98*, pages 52–62, 1998.

[18] G. Halin and D. Hanser. Vers une visualisation contextuelle de la conception coopérative. In *IHM 2003: Proceedings of the 15th French-speaking conference on human-computer interaction on 15eme Conference Francophone sur l'Interaction Homme-Machine*, pages 208–211, New York, NY, USA, 2003. ACM Press.

[19] J. Hathaway. Project juggling: the evolution of a process. In *Proceedings of the 32nd annual ACM SIGUCCS conference on User services*, pages 247–249, 2004.

[20] S. Havre, E. Hetzler, K. Perrine, E. Jurrus, and N. Miller. Interactive visualization of multiple query results. In *InfoVis 2001*, pages 105–112, 2001.

[21] I. Herman, G. Melançon, and M. S. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, Jan. 2000.

[22] B. Hetzler, P. Whitney, L. Martucci, and J. Thomas. Multi-faceted insight through interoperable visual information analysis paradigms. In *InfoVis 1998*, pages 137–144, 161, 1998.

[23] T. Igarashi, S. Matsuoka, and H. Tanaka. Teddy: a sketching interface for 3D freeform design. In *SIGGRAPH '99*, pages 409–416, 1999.

[24] B. S. Johnson. *Treemaps: Visualizing Hierarchical and Categorical Data*. PhD thesis, University of Maryland, 1993.

[25] G. Jones. One solution for project management. In *Proceedings of the 29th annual ACM SIGUCCS conference on User services*, pages 65–69, 2001.

[26] D. Joslin and W. Poole. Agent-based simulation for software project planning. In *Proceedings of the 37th conference on Winter simulation*, pages 1059–1066, 2005.

[27] D. A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, Jan. 2002.

[28] D. A. Keim, F. Mansmann, J. Schneidewind, and T. Schreck. Monitoring network traffic with radial traffic analyzer. In *2006 IEEE Symposium On Visual Analytics Science And Technology (VAST)*, pages 123–128, 2006.

[29] D. A. Keim, J. Schneidewind, and M. Sips. CircleView: a new approach for visualizing time-related multidimensional data sets. In *AVI '04: Proceedings of the working conference on Advanced visual interfaces*, pages 179–182, New York, NY, USA, 2004. ACM Press.

[30] J. Kontio, J.-P. Jokinin, and E. Rosendahl. Visualizing and formalizing risk information: an experiment. In *Proceedings of the 10th International Symposium on Software Metrics*, pages 196–206, 2004.

[31] G. Kumar and M. Garland. Visual exploration of complex time-varying graphs. In *InfoVis 2006*, pages 805–812, 2006.

[32] H. Lee-Kwang and J. Farrel. The ssd graph: a tool for project scheduling and visualization. *IEEE Transactions on Engineering Management*, 35(1):25–30, Feb. 1988.

[33] P. Lempp. A possibility to make project progress visible: the progress diagram within the epos system. *ACM SIGSOFT Software Engineering Notes*, 11(2):32–35, Apr. 1986.

[34] Y. Livnat, J. Agutter, S. Moon, and S. Foresti. Visual correlation for situational awareness. In *InfoVis 2005*, pages 95–102, 2005.

[35] M. Mantei. The effect of programming team structures on programming tasks. *Communications of the ACM*, 24(3):106–113, Mar. 1981.

[36] B. Meyer. The software knowledge base. In *Proceedings of the 8th international conference on Software engineering*, pages 158–165, 1985.

[37] Microsoft Corporation. Microsoft Project. http://office.microsoft.com/project/. Accessed 15 May 2007.

[38] Microsoft Corporation. Microsoft Visio. http://office.microsoft.com/visio/. Accessed 15 May 2007.

[39] Z. Mihajlovic and D. Velasevic. Tracking software projects with the integrated version control in smit. *ACM SIGSOFT Software Engineering Notes*, 26(2):38–43, Apr. 2001.

[40] M. J. Muller, J. G. Smith, J. Z. Shoher, and H. Goldberg. Privacy, anonymity and interpersonal competition issues identified during participatory design of project management groupware. In *Proceedings of the 6th annual ACM SIGUCCS conference on User services*, pages 82–87, 1990.

[41] F. Nightingale. *Notes on matters affecting the health, efficiency, and hospital administration of the British army : founded chiefly on the experience of the late war*. London: Harrison and Sons, 1858.

[42] D. A. Norman. *The Design of Everyday Things*, page 209. Basic Books, 2002 edition, 2002.

[43] M. L. Northway. *A Primer of Sociometry*. University of Toronto Press, 2nd edition, 1967.

[44] J. P. Peddicord. A.p.i.s. academic project information system. In *Proceedings of the 6th annual ACM SIGUCCS conference on User services*, pages 74–81, 1978.

[45] W. Playfair. *The statistical breviary; shewing, on a principle entirely new, the resources of every state and kingdom in Europe; illustrated with stained copper plate charts, representing the physical powers of each distinct nation with ease and perspicuity. To which is added, a similar exhibition of the ruling powers of Hindoostan*. London: for J. Wallis, 1801.

[46] S. Polyviou, G. Samaras, and P. Evripidou. A relationally complete visual query language for heterogeneous data sources and pervasive querying. In *ICDE 2005. Proceedings of 21st International Conference on Data Engineering*, pages 471–482, 2005.

[47] S. F. Roth, P. Lucas, J. A. Senn, C. C. Gomberg, M. B. Burks, P. J. Stroffolino, J. A. Kolojejchick, and C. Dunmire. Visage: A user interface environment for exploring information. In *InfoVis 1996*, pages 3–16, 1996.

[48] S. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*, page 46. Prentice Hall, first edition, 1995.

[49] C. E. Scheidegger, H. T. Vo, D. Koop, J. Freire, and C. T. Silva. Querying and creating visualizations by analogy. *IEEE Transactions on Visualization and Computer Graphics*, 13(3):1560–1567, 2007.

[50] B. Shneiderman. The eyes have it: a task by data type taxonomy for information visualizations. In *IEEE Symposium on Visual Languages*, pages 336–343, 1996.

[51] V. Sinha and D. R. Karger. Magnet: Supporting navigation in semistructured data environments. In *SIGMOD*, pages 97–106, 2005.

[52] I. Spence. No humble pie: The origins and usage of a statistical chart. *Journal of Educational and Behavioral Statistics*, 30(4):353–368, 2005.

[53] J. Stamey and T. Honeycutt. Quad charts in software project management. *Journal of Computing Sciences in Colleges*, 21(2):322–327, Dec. 2005.

[54] J. Stasko and E. Zhang. Focus+context display and navigation techniques for enhancing radial, space-filling hierarchy visualizations. In *InfoVis 2000*, pages 57–65, 2000.

[55] C. Stolte, D. Tang, , and P. Hanrahan. Polaris: A System for Query, Analysis, and Visualization of Multidimensional Relational Databases. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):52–65, 2002.

[56] M. Stonebraker. Visionary: A next generation visualization system for data bases. In *SIGMOD*, page 635, 2003.

[57] P. Szekely, M. Becker, S. Fitzpatrick, G. Gati, D. Hanak, J. Jin, G. Karsai, R. T. Maheswaran, B. Neches, C. M. Rogers, R. Sanchez, and C. van Buskirk. Csc: Criticality-sensitive coordination. In *Proceedings of the fifth international joint conference on Autonomous agents and multiagent systems*, pages 1441–1442, 2006.

[58] A. Trigoni. Interactive query formulation in semistructured databases. In *5th International Conference on Flexible Query Answering Systems*, pages 356–369, 2002.

[59] W. Wang and F. Lillehagen. A cooperative hypermedia solution to work management in real-time enterprises. In *Proceedings of the fourteenth ACM conference on Hypertext and hypermedia*, pages 196–197, 2003.

[60] M. O. Ward. XmdvTool: Integrating multiple methods for visualzing multivariate data. In *IEEE Visualization*, pages 326–333, 1994.

A Holistic Framework for Hand Gestures Design

Juan Wachs[1], Helman Stern[2] and Yael Edan[2]

[1]*Department of Computer Science*
Naval Postgraduate School, 700 Dyer Road
Monterey, CA, 93943-5001, USA
jpwachs@nps.edu
[2]*Department of Industrial Engineering and Management,*
Ben-Gurion University of the Negev, Beer Sheva, 84105, Israel
{helman,yael}@bgu.ac.il

Abstract

Hand gesture based interfaces are a proliferating area for immersive and augmented reality systems due to the rich interaction provided by this type of modality. Even though proper design of such interfaces requires accurate recognition, usability, ergonomic design and comfort. In most of the interfaces being developed the primary focus is on accurate gesture recognition.

Formally, an optimal hand gesture vocabulary (GV), can be defined as a set of gesture-command associations, such that the time τ to perform a task is minimized over all possible hand gestures in our ontology. In this work, we consider three different cost functions as proxies to task completion time: intuitiveness $Z_1(GV)$, comfort $Z_2(GV)$ and recognition accuracy $Z_3(GV)$. Hence, we can establish that $Max(Z_{i)}(GV)$: i=1,2,3) over all GV's is our multiobjective problem (MOP).

Because finding the solutions to the MOP requires a large amount of computation time, an analytical methodology is proposed in which the MOP is converted to a dual priority objective problem where recognition accuracy is considered of prime importance, and the human performance objectives are secondary.

This work, as opposed to previous research done by the authors, is focused on two aspects: First,a modified cost function for an enhanced simulated annealing approach is explained and implementation issues are discussed. Second, a comparative study is performed between hand gesture vocabularies obtained using the methodology suggested, and vocabularies hand picked by individuals.. The superiority of our method is demonstrated in the context of a robotic vehicle control task using hand gestures.

1. Introduction

The majority of human machine interfaces for everyday device control aims for affordable prices while mimicking realistic natural interactions. This type of interface can be activated by voice, face, hand and body posture recognition algorithms. Most of them are designed to achieve a high recognition performance while allowing the user to interact with the systems similar to interactions with another human. Human-robot interaction was exploited in [1] in the context of ambient intelligence (intelligence algorithms involving measurement, transmission, modeling, and control of environmental information) for human detection and gesture recognition. Hand detection and pose recognition was achieved in [2] through an infra-red time-of-flight range camera. The author's interface system was able to recognize 7 DoFs of a human hand with a 2-3 Hz frame rate. In [3] a method for recognizing hand gestures using depth image data acquired from active vision hardware was suggested. The authors are able to recognize different static poses while tracking the hand in real time. The authors were motivated by the development of an interface to control home appliances. A notable work in home appliance control was done in [4]. The authors propose an universal remote control system based merely on hand gestures. In this system, the user first selects the device to be controlled by pointing it with his hand. Then, the user operates it through 10 predefined basic hand motions. Marcel [5] developed a system that combined face tracking with hand gestures recognition based on face location and body anthropometry. This system was capable of recognizing a five gesture vocabulary in uniform and cluttered environments, however no applications were suggested for such an interface. Dynamic gesture recognition to drive mobile phone applications was developed by [6] based on accelerometers attached to the mobile phone. The authors present a proof of concept of their system through a "navigate and select" application, such as Google Earth. Wireless communication is also used by [7] for voice and real-time continuous sign language recognition. The authors implemented their system in a post-wearable PC in the domain of ubiquitous computing applications which allowed the users to freely move with their portable terminal while naturally interacting with the embedded-ubiquitous environment. A hybrid system capable of using information from faces and voices to recognize people's emotions was developed in the PHYSTA project [8]. In [9] a system was developed capable to incrementally learn to recognize affective states from body postures for human-robot interaction.

Two types of hand gesture interfaces have been distinguished in human-machine interaction according to their objective: the first is designed to cope with the challenge of hand gesture recognition with a high accuracy and speed, and the other is focused on the ergonomic aspects of the hand gesture vocabulary design. In all the research presented above enormous efforts were invested in achieving the first objective, (technical focused), but the second objective was not addressed. Understanding the user's physiologic and cognitive needs is one of the key tasks associated with an efficient and natural hand gesture based interface. To tackle that task, machine vision and analysis techniques have to be developed, while, at the same time, psychological and linguistic analyses of hand gestures must be considered. An example of the second type of hand gesture interfaces can be found in [10], where intuitive hand gestures are selected in a fashion that allows the user to act more naturally since no cognitive effort is required in mapping function keys to robotic hand actions. This system, like others, is based in navigation control. The author's selection of gestures can be criticized since their interpretation of an intuitive gesture-command association may not suit others' cognitive perception of intuitiveness. This issue was addressed in [11] where it was found that people consistently used the same gestures for specific commands. In particular they found that people are also very proficient at learning new arbitrary gestures. In [12] it was found that test subjects used very similar gestures for the same operations. All these may indicate that there may be intuitive, common principles in gesture communication. A notable work discussing an ergonomic based approach for hand gestures design is presented in [13], where the comfort associated with specific gestures when they are performed rapidly and repeatedly is considered. The authors conclude that designers of gesture languages for computer input should minimize the use of those hand gestures associated with upper extremity discomfort. Also in [14] a similar conclusion is achieved. The authors used a biomechanics based objective function to reflect comfort in the framework of a hand gesture based interface.

Previously, in [15] and [16] we have shown a methodology for the design of a gesture vocabulary that is both intuitive and comfortable on the one hand, and can be recognized with high accuracy, on the other. A two-step procedure for solving the gesture vocabulary design problem was introduced. This procedure was formulated as a multiobjective optimization problem (MOP). The first step is to decide on a task-dependent set of commands to be included in the vocabulary such as; "move left", "increase speed", etc. The second step is to decide how to express the command in gesture form i.e., what physical expression to use, such as waving the hand left to right or making a "V" sign with the first two fingers. The association (matching) of each command to a gesture expression is defined here as a "gesture vocabulary" (GV). In this paper, the gesture-command matching algorithm based on simulated annealing is discussed for the first time, and new experiments comparing "human hand picked" vocabularies with automated hand gesture vocabulary design are presented.

In the next section the GV design problem is defined. This is followed in section 3 by a description of the main methodology; comprised of hand gesture factor determination, gesture subset selection, command-gesture matching, and selection of pareto optimal multiobjective solutions. In section 4, the extended simulated annealing approach to solve the optimal gesture-command association is presented. Section 5 compared our automated approach with human selected GVs. Section 5 provides conclusions.

2. Problem Statement

A suitable definition of a hand gesture vocabulary (GV) is the set of gesture-command pairs that minimizes the time τ required for a user/s to perform a task/s. The number of commands is fixed, and determined by the given task. The set of gestures G_n is obtained from a large set of postures; a "master-set" of gestures, denoted by G_m.
Three performance measures are used as proxies for the task completion time τ. intuitiveness $Z_1(GV)$, comfort $Z_2(GV)$ and recognition accuracy $Z_3(GV)$. The first two measures are related to ergonomic side, while the last is strictly related to the technological aspect.

The problem is to find a GV that maximizes the proxies, achieving a minimal performance time, over all feasible gesture vocabularies, Γ. This multi-objective problem (MOP) is complex given that the performance time is not a well-behaved function on the proxies manifolds. Moreover, there exist conflicting solutions where all the objectives cannot be maximized simultaneously. This can be overcome by allowing the decision maker to select the best GV according to his own preferences.

$$Max\ Z_1(GV), Max\ Z_2(GV), Max\ Z_3(GV) \qquad (1)$$
$$GV \in \Gamma$$

Let us define Z1, the intuitiveness, of the GV as the naturalness of expressing a given command with a gesture. We recognize two types of intuitiveness: direct and complementary.

Let p be defined as an assignment function where p(i)=j indicates that the command i is assigned to gesture j. Consequently, the direct intuitiveness, $a_{i,p(i)}$ is expressed by the strength of the association between command i and its matched gesture p(i). Following the same concept, complementary intuitiveness, $a_{i,p(i),j,p(j)}$ is the level of association expressed by the matching of complementary gesture pairs (p(i), p(j)) to complementary command pairs (i,j). The total intuitiveness is shown in (2).

$$Z_1(GV) = \sum_{i=1}^{n} a_{i,p(i)} + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{i,p(i),j,p(j)} \quad (2)$$

Let us define Z_2 as the Stress/Comfort needed to perform a gesture. Obviously, there are gestures that are easier to perform than others. Total stress is a scalar value equal to the sum of the individual stress values to hold the postures, and to perform transitions between them, weighted by duration and frequency of use.

Thus s_{kl} is the physical difficulty of a transition between gestures k and l. The duration to reconfigure the hand between gestures k and l is represented by d_{kl}. The symbol f_{ij} stands for the frequency of transition between commands i and j. The value K is a constant and is used to convert stress into its inverse measure comfort.

$$Z_2(GV) = K - \sum_{i=1}^{n}\sum_{j=1}^{n} f_{ij}\, d_{p(i)p(j)} s_{p(i),p(j)} \quad (3)$$

Accuracy is a measure of how well a set of gestures can be recognized. To obtain an estimate of gesture accuracy, it is necessary to train a gesture recognition system on a set of sample gestures for each gesture in G_n. The number of gestures classified correctly and misclassified is denoted as Tg and Te, respectively. The gesture recognition accuracy is denoted by (4).

$$Z_3(GV) = \left[\left(T_g - T_e\right)/T_g\right]100 \quad (4)$$

3. Main Methodology

One approach to solve (1) is find the performance measures over the set of all feasible GVs (complete enumeration). This approach is untenable, for even reasonable size vocabularies, thus, a dual priority objective optimization, (where recognition accuracy is considered of prime importance, and the human performance objectives are secondary) is proposed as a more tractable approach. Lets combine the intuitive and comfort objectives into one objective $\overline{Z}$ using weights w_1 w_2, and let A_{min}, be the minimum acceptable accuracy. Then we obtain:

$$Max\overline{Z}(GV) = w_1 Z_1(GV) + w_2 Z_2(GV) \quad (5)$$

$$GV \in \Gamma$$

$$s.t. \quad Z_3(GV) \geq A_{min} \quad (6)$$

The architecture of the solution methodology is comprised of four modules (Fig. 1). In Module 1 human psycho-physiological input factors are determined. In Module 2 gesture subsets, satisfying (6) are determined; Module 3 constitutes a command - gesture matching procedure. Finally, the set of Pareto optimal solutions is found in Module 4.

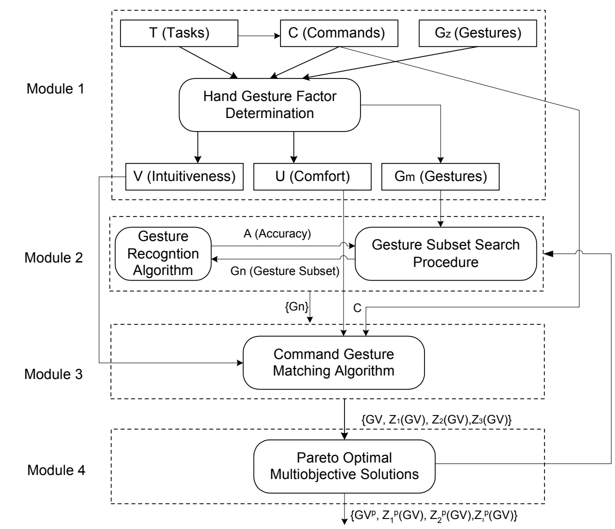


Figure 1. Architecture of optimal hand gesture vocabulary solution procedure

3.1. Module 1: Hand Gesture Factor Determination

The input parameters to the Module 1is the task set T, a large gesture master set G_z and the set of commands C. The procedure to obtain the intuitiveness V, comfort U, and gesture G_m, matrices is explained in [14]. For each task t_i , a set C of c_i commands are defined, as the union of all the task commands. Given the sequence of commands needed to complete a task the command transition matrix (F) is computed. The f_{ij} entries in F, represent the frequency that a command c_j is evoked given that the last command was c_i.

Since the set of all possible gestures is infinite, we established a set of plausible gesture configurations based on an articulated model including finger positions (extended, spread), palm orientations (up, down sideways), and wrist rotations (left, middle, right) as the primitives, see Figure 2. Moreover, the gesture set is further reduced by considering the normalized popularity of gesture among the users. This final set is called the Gesture Master Set (G_m).

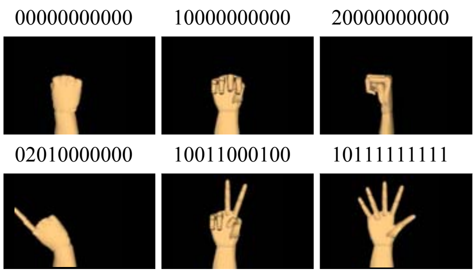


Figure 2. Articulated hand gesture model

Once the gesture set is reduced, the intuitiveness matrix I can be obtained. The entries of this matrix a_{ik} represent the naturalness of using gesture i for command k. In the same fashion, the complementary intuitive matrix (I') is attained, where the entry a_{ijkl} express the naturalness of matching up a pair of complementary commands (i, j) with a pair of complementary gestures (k,l). Denote V=[I, I'] as the set of matrices including both the direct and complementary matrices.

The fatigue (or comfort) indices are arranged in a matrix (S) where the element s_{ij} represents the physical difficulty of performing a transition from gesture i to gesture j. An entry u_{ijkl} in the comfort matrix (U) is defined as K-f_{ij} x s_{kl} where the last term represents the frequency of transition between commands i to j times the stress of a command transition k to l given that i and j are paired with gestures k and l, respectively.

3.2. Module 2: Gesture Subset Selection

The inputs of Module 2 are the reduced master set of gestures G_m, and a recognition algorithm to determine A. An iterative search procedure to find a set of gesture subsets $\{G_n\}$ is used in this module, to satisfy a given accuracy (6). The subset search procedure is based on the properties of the confusion matrix of the multi gesture recognition algorithm, and is called: Confusion Matrix Derived Solution Method (CMD) [14].

The CMD method consists of three steps: (i) train the recognition algorithm for the gestures in G_m , and let C_m be the resulting confusion matrix. The confusion matrix is obtained directly from the partition result of the training set using a supervised FCM optimization procedure, [17], (ii) find a submatrix C_n from C_m such that the recognition accuracy is highest, equal or below A_{min} (6), and (iii) repeat (ii) until a given number of solutions are found. The CMD algorithm obtains $\mathcal{N}$ solutions (or all the solutions with associated accuracy above a given minimum allowed A_{min} if less than $|\mathcal{N}|$). Each iteration of the CMD algorithm generates a new solution by excluding each time a different gesture, from the subset of gestures of the current solution, and adding a new gesture from the master set. The number of solutions $|\mathcal{N}|$, is determined by the number of GV's that we want to consider based on the three measures Z_1, Z_2 and Z_3. This is usually specified by the decision maker.

3.3. Module 3: Command-Gesture Matching

The inputs to the third module are the intuitiveness V and comfort U matrices, the command set C, and the subset of gestures G_n. The purpose of this module is to match the set of gestures G_n to the set of commands, C, such that the human measures are maximized. The resulting gesture-command assignment constitutes a gesture vocabulary, GV.

Given a single set of gestures $G_n \in \mathcal{N}$ found from module 2, the gesture-command matching can be represented as a quadratic integer assignment problem (QAP) [8] and is formulated in (7)-(10).

$$\max \overline{Z}(G_n^*) = w_2 \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} u_{ijkl} x_{ik} x_{jl} + w_1 \left[\sum_{i}^{n}\sum_{j}^{n} v_{ij} x_{ij} + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} v_{ijkl} x_{ik} x_{jl} \right] \quad (7)$$

$$\sum_{j=1}^{n} x_{ij} = 1, \quad i = 1,..,n, \quad (8)$$

$$\sum_{i=1}^{n} x_{ij} = 1, \quad j = 1,...,n, \quad (9)$$

$$x_{ij} \in \{0,1\}; \quad i = 1,...,n, \quad j = 1,...,n, \quad (10)$$

Let x_{ij} be the binary assignment variable. x_{ij} is equal to 1 if command i is assigned to gesture j, and zero otherwise. Equation (8) constraints each command to be matched with exactly one gesture. Equation (9) constrains each gesture to be matched with exactly one command. An enhanced simulated annealing is adopted to solve the QAP and it will be described in Section 4.

For each subset G_n found on Module 2, the QAP is solved by varying the weights such that w1+w2=10. This results in a set of GV solutions corresponding to each G_n in $\mathcal{N}$.

3.4 Module 4: Pareto Optimal Multiobjective Solution

Each of the $\mathcal{N}$ solutions (gesture subsets G_n) from Module 2, can result in $\mathcal{M}$ derived solutions. Each combination of the weights, for a given G_n, results in a new solution GV.

Thus a total of $\mathcal{N}$x$\mathcal{M}$ candidate GV's solutions are expected. Each of these solutions may be represented as a point in 3D space, (Z_1,Z_2,Z_3). The total set of multiobjective candidate solutions is then $\{Z_1(GV), Z_2(GV), Z_3(GV): GV=\{1,\ldots, \mathcal{N}x\mathcal{M}\}$.

A set of Pareto solutions exists for this 3D manifold surface. A Pareto solution is one that is not dominated by any other solution. That is, a Pareto solution is one in which one cannot increase one performance measure without decreasing at least one of the others. The Pareto solutions offer a reduced set of candidate solutions from which a decision maker can select the GV that meets his/her internal preferences.

4. Solving the QAP by annealing

A simulated annealing scheme is used to solve the QAP since it provides improved solutions for several of the largest combinatorial problems available in literature and requires low computational effort [18]. The core idea of this approach is defining a "smart" strategy to find certain uphill steps to avoid convergence in local minima. This means: (a) move from the current solution to a neighboring one efficiently, (b) compute the change in the objective function, (c) if the objective function is improved by the step accept it, otherwise (d) accept the step with a probability P(accept)= $e^{-\delta/kT}$.

Where δ is the perturbed solution, T is a value representing the absolute temperature (in the analogy used to simulate energy levels in cooling solids) and k is Boltzmann's constant.

4.1. Annealing formulation for the QAP

The objective solution presented in [16] is given by:

$$\min \sum_{i=1}^{n}\sum_{a=1}^{n} B_{ik}x_{ik} + \sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} C_{ijkl}x_{ik}x_{jl} \tag{11}$$

$$\sum_{j=1}^{n} x_{ik} = 1, \quad k = 1,..,n, \tag{12}$$

$$\sum_{j=1}^{n} x_{jl} = 1, \quad l = 1,...,n, \tag{13}$$

$$x_{ik} \in \{0,1\}; \quad i = 1,...,n, \quad k = 1,...,n, \tag{14}$$

Where B_{ik} is the cost of assigning facility i to location k and the cost by the double assignment of i to k and j to l is represented by C_{ijkl}. This value can also be seen as the flow F_{ij} between facilities i and j times the distance D_{kl} between locations k and l.

For notation simplicity we will denote any feasible solution by a permutation p of the integers from 1 to n where $p(i)$ represents the chosen location for facility i.

Simulated annealing (SA) starts the searching process from a random permutation of the facilities. A neighborhood move is achieved by exchanging the pair of facilities i and j and evaluating the relative change in the objective function using the formula:

$$\begin{aligned}\delta &= B_{ip(j)} + B_{jp(i)} - B_{ip(i)} - B_{jp(j)} \\ &+ 2\sum_{h\neq i,j}\left[(F_{jh} - F_{ih})(D_{p(i)p(h)} - D_{p(j)p(h)})\right]\end{aligned} \tag{15}$$

The moves that improve the objective function (11) (i.e. $\delta \le 0$) are accepted while uphill steps ($\delta \ge 0$) are accepted with a probability P(accept)= $e^{-\delta/T}$ by drawing a random number x from a uniform distribution [0,1] and accepting the exchange if $x \le e^{-\delta/T}$. In our scheme, we try to maximize our cost function given by (7). In this context, a neighborhood move consists of exchanging two commands i and j and the relative change in the objective function. The marginal change δ is obtained by the contribution obtained by the exchange of the pair of gestures, minus the value associated to the current matching. When other command-gesture associations are affected by the exchange, the marginal change must be calculated for each of the remaining associations with respect to each of the pair exchanged.

Let δ_I, δ_S, δ_{IC} denote the particular contribution of the exchange of commands on the intuitiveness, the stress and the complementary intuitiveness respectively.

Let h_i be the scaling factor for each the intuitiveness, the stress and the complementary intuitiveness.

Let k_j be the weights assigned by the decision maker reflecting the importance of each term.

Let $\eta(i,j)$ be a function that is equal to one if the commands i,j are complementary, otherwise zero.

$$\delta_I = h_1 k_1 (a_{r,p(s)} + a_{s,p(r)} - a_{r,p(r)} - a_{s,p(s)}) \tag{16}$$

$$\begin{aligned}\delta_S &= -h_2 k_2 ((s_{r,r} - s_{s,s})(f_{p(s),p(s)} d_{p(s),p(s)} - f_{p(r),p(r)} d_{p(r),p(r)}) + \\ &(s_{r,s} - s_{s,r})(f_{p(s),p(r)} d_{p(s),p(r)} - f_{p(r),p(s)} d_{p(r),p(s)}) + \\ &\sum_{k\neq r,s}(s_{k,r} - s_{k,s})(f_{p(k),p(s)} d_{p(k),p(s)} - f_{p(k),p(r)} d_{p(k),p(r)}) + \\ &(s_{r,k} - s_{s,k})(f_{p(s),p(k)} d_{p(s),p(k)} - f_{p(r),p(k)} d_{p(r),p(k)})\end{aligned} \tag{17}$$

$$\begin{aligned}\delta_{IC} &= 2h_3 k_3 (\eta(r,s) k_1 (a_{s,r,p(r),p(s)} - a_{s,r,p(s),p(r)}) \\ &\sum_{k\neq r,s}(\eta(k,s)(a_{k,s,p(k),p(r)} - a_{k,s,p(k),p(s)}) + \\ &\eta(k,r)(a_{k,r,p(k),p(s)} - a_{k,r,p(k),p(r)})))\end{aligned} \tag{18}$$

Hence the relative change is evaluated as $\delta=\delta_I+\delta_S+\delta_{IC}$. In our application we used $h_1= h_3=1$ and $h_2=0.001$.

4.2. Neighborhood structure

The next potential solution for our particular neighborhood structure is chosen based on a pseudo-random method, such that the pair exchange follows the sequence:

(1,2),(1,3),……,(1,n),(2,3),…,(n-1,n),
(1,2)……

The exchange is accepted according to the result of δ as described in Section 8.1. The probability of exchange is regulated by the temperature, which in turn drops after each attempted pair exchange. The temperature changes between initial and final values T_0 and T_f, respectively, according to (19):

$$T_{n+1} = T_n/(1+\beta T_n), \quad where\ \beta << T_0 \tag{19}$$

The cooling scheme is controlled by specifying a number of steps M using (20).

$$\beta = (T_0 - T_f)/MT_0T_f \tag{20}$$

Four trials were performed in the gesture-command matching problem where M was 6,000.

T_0 and T_f can be obtained using (19) by finding the maximum and minimum positive values of δ when running the neighborhood search for 1000 iterations.

$$\begin{aligned} T_0 &= \delta_{min} + \tfrac{1}{10}(\delta_{max} - \delta_{min}) \\ T_f &= \delta_{min} \end{aligned} \tag{21}$$

The optimal temperature is the one that corresponds to the exchange of gestures that yields the maximum in (7). To avoid being trapped in a local maximum when consecutive marginal changes δ are rejected, we proceed in the following manner:

a) Accept the next negative contribution δ.
b) Set the optimal temperature to the current one.
c) Cooling is stopped (β=0).

This procedure was implemented for the optimal command-gesture matching problem, and we found that all the results obtained for 15 problems of a robotic arm example were global optima. Moreover, this approach was originally applied by [16] using the contribution δ as expressed in (15) and they reported solutions within 1% of the best known solutions for *n*=50 and 100.

5. Experiments and Results

A robotic vehicle control task using hand gestures is used to test the procedure explained in the previous chapters.

5.1. The Pareto Set of Solutions

Eight 'navigational' (directional) commands to control the direction of movement of the robot were chosen. From a master set of 22 postures, sets of 8 gestures are extracted and matched to the 8 commands (see Fig 3). The commands used were: start, finish, left, right, forward, backward, fast and slow.

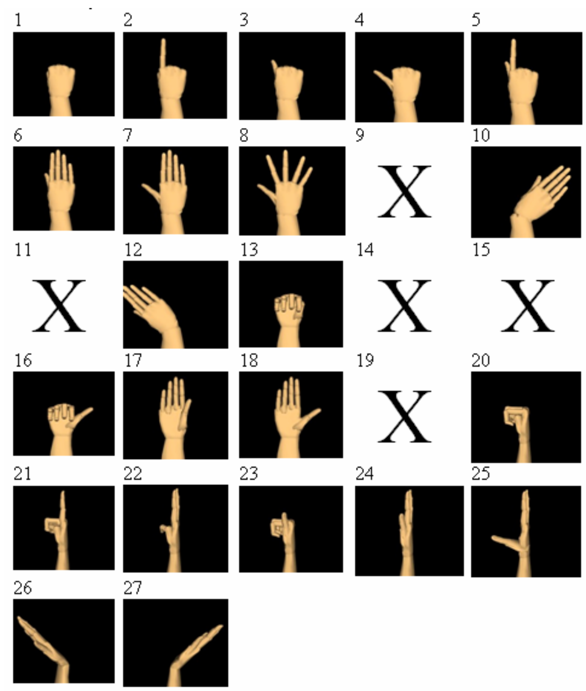


Figure 3. Gesture master set and command set for the robotic vehicle task

The algorithm generated eight solutions, where the minimal acceptable accuracy was set to 96.25 percent. Each of these solutions produced a set of 11 GV candidates. (a total of 88 GV's from eight different subsets of gestures G_n, and 11 weight combinations). The plots in Fig. 4 show the intuitiveness versus comfort trade offs for each G_n and its associated accuracy $A(G_n)$. Associated GV solutions are connected together, forming a curve for a given G_n. This family of curves is shown in a space orthogonal to the recognition accuracy coordinate (Figure 4). From this set of solutions a Pareto set of 8 GV's was obtained.

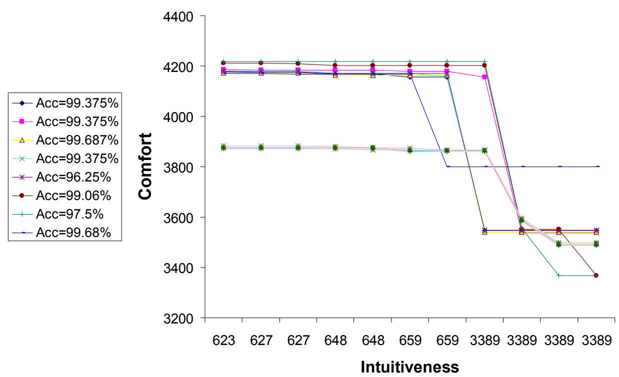


Figure 4. Intuitiveness vs. comfort families of 8 curves

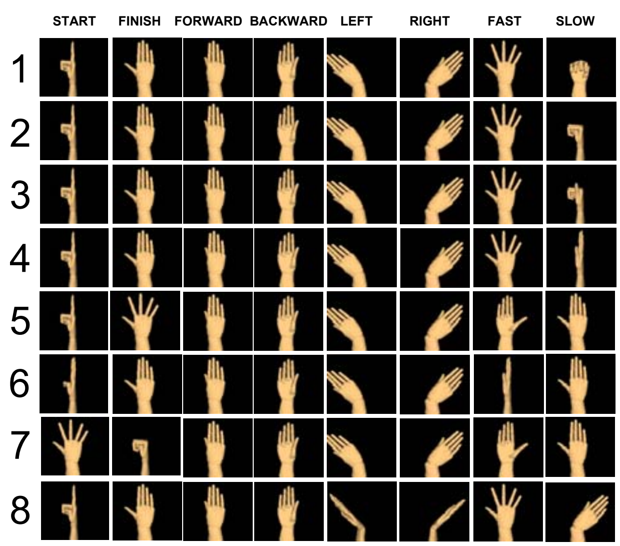


Figure 5. Pareto front GV solutions

5.2. Human Vs Computer Hand GV selection

In this section we aim to determine whether the automated methodology is better than a "hand-picked" method according to ergonomic and technical parameters. This issue was addressed through two small experiments in the context of a robotic arm "pick & place" task.

In the experiment used to obtain the natural association between commands and gestures the user was presented with a sequence of commands required to perform the "pick and place" task. The user manipulated a hand model until it was configured to represent the desired gesture, matching the displayed command. One by one all the commands were presented and their respective gesture matched according to the user desires. Once this data was collected for 35 users, two experiments were conducted.

In the first one, we investigated whether the automated system could find better associations than those provided by subjective experiments. Given a GV selected by a user, is it possible to find better gesture-command associations?

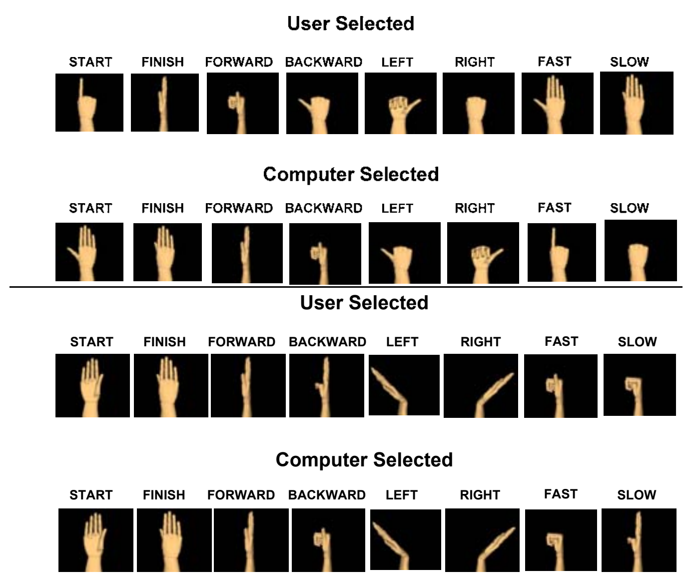
We used the results of the intuitiveness experiment to extract eight GV's from eight different users out of a set of 35 users. We selected those users that selected gestures that belonged to the reduced gesture set. We supplied each G_{n1} to the automated system and obtained new gesture-command associations. The results are compiled in Table 1.

Table 1. Human Vs Computer Hand GV selection

GV	Comfort (Z1)		Intuiteveness (Z2)		Accuracy (Z3)
	Human	Comp.	Human	Comp.	
1	3625	3625	2960	2960	95.30%
2	3661	**3807**	24	**3296**	87.50%
3	3617	3617	2706	2706	97.10%
4	3621	**3626**	**2854**	2851	99.00%
5	3569	3569	3488	3488	91.80%
6	3628	**3631**	2697	**2973**	93.70%
7	3615	3615	2524	2524	90.90%
8	3683	**3815**	552	**3334**	91.50%

The "Comfort" and "Intuitiveness" measures of eight GVs were obtained using a subjective test (Human) and the automated method (Comp). There is only one column for the Accuracy measure since both comparisons were based on the same subset of gestures, and the recognition accuracy is only a function of the gestures used, and not of their associations.

In all the GVs compared, the automated method performed better (GV2, GV6 and GV8) (see Figure 1) or equal (GV1, GV3, GV5 and GV7). There was only one case (GV4) where the GV selected by the user was more intuitive than the one selected by the automated approach; however this GV was lower in comfort.



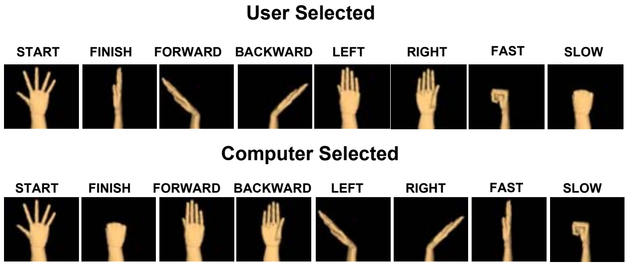


Figure 6. Three dominating solutions GV2, GV6 and GV8.

The GV's generated automatically differed from the human selected ones by at least three gesture-command matchings (GV 6) and at most eight gesture-command matchings (GV 2).

In the second experiment, we compared eight GVs obtained from the Pareto front, from the generated solutions by our methodology, to the same eight GVs created by eight users tested in the previous experiment.

The results are summarized in Table 2(a) and (b).

Table 2. Solutions found on the Pareto frontier

GV	Comfort Z1	Intuiteveness Z2	Accuracy Z3
1'	3546	3389	99.38%
2'	3549	3383	99.38%
3'	3548	3380	96.25%
4'	3552	3376	99.06%
5'	3541	3157	99.69%
6'	3556	3151	97.50%
7'	3539	3142	99.38%
8'	3801	3020	99.69%

Note that all the solutions found through the automated procedure are superior to those suggested by the user in two measures out of three (at least): Accuracy and Intuitiveness; except for two solutions GV3' and GV6', which were significantly less intuitive. The eight solutions are presented in Figure 5.

Both examples presented in this chapter show that hand gesture vocabularies obtained by the automated system have higher or equal ergonomic and technical measures than those proposed by the user in most cases.

6. Conclusions

Proper design of hand gesture-based human-machine interfaces requires accurate recognition, ergonomic design and comfort. Unfortunately, in most interfaces developed, efforts are focused primarily in accurate recognition of the gestures, which is a technical consideration only. In this work, we considered three different cost functions as proxies to task completion time: intuitiveness $Z_1(GV)$, comfort $Z_2(GV)$ and recognition accuracy $Z_3(GV)$. We established that the set of optimal hand gesture vocabularies can be accurately formulated as a maximization of the individual measures (Z1,Z2 and Z3) in a multiobjective problem fashion. The solutions are obtained by the Pareto points and the final solutions are chosen by the decision maker according to his preferences over the three objectives.

Associating a subset of gestures with commands was presented as a binary integer quadratic assignment problem which was solved by simulated annealing. The first contribution of this work is to present the modified cost function for the enhanced simulated annealing. The second contribution was a comparative study between hand gestures vocabularies obtained using the methodology suggested, and vocabularies obtained by user hand selections.

Two experiments were carried out to show the superiority of our method. In the first one we showed that the automated system can find better or equal associations than those provided from subjective experiments (using the same subset of gestures). In the second experiment, we compared the previous eight GVs selected by the user to those obtained through the Pareto front, from the generated solutions by our methodology. All the solutions found through the automated procedure were superior to those suggested by the user in two measures out of three (at least). These results indicate, in a quantitative fashion, the importance of considering technical and ergonomic aspects for a successful development and design of hand gesture interface systems.

7. Acknowledgements

This research was partially supported by the Paul Ivanier Center for Robotics Research & Production Management at Ben-Gurion University of the Negev.

8. References

[1] Kubota, N. Tomioka, Y. "Evolutionary robot vision for human tracking of partner robots in ambient intelligence" *IEEE Congress on Evolutionary Computation,* 2007. pp.1491-1496.

[2] Breuer, P. Eckes, C, Müller, S. "Hand Gesture Recognition with a Novel IR Time-of-Flight Range Camera: A Pilot Study." In *Proc. of International Conference on Computer Vision/Computer Graphics Collaboration Techniques*. Lecture Notes in Computer Science. Berlin: Springer, 2007. pp. 247-260.

[3] Xia Liu Fujimura, K. "Hand gesture recognition using depth data." in *Proceedings of IEEE International Conference on Automatic Face and Gesture Recognition,* 2004. pp. 529- 534.

[4] J.H. Do, H. Jang, S. H. Jung, J. J. Z. Bien. "Soft remote control system in the intelligent sweet home" in IEEE Conference on Intelligent Robots and Systems, 2005. pp. 3984- 3989.

[5] Marcel, S.; Bernier, O.; Viallet, J.-E.; Collobert, D. "Hand gesture recognition using input-output hidden Markov models." in *Proceedings of Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000. pp. 456 - 461

[6] Majoe, D. Schubiger, S. Clay, A. Arisona, S.M. "SQEAK: A Mobile Multi Platform Phone and Networks Gesture Sensor." *In the 2nd International Conference on Pervasive Computing and Applications, 2007. ICPCA 2007.* pp. 699-704

[7] J.H. Kim and K.-S. Hong. "Multi-Modal Recognition System Integrating Fuzzy Logic-based Embedded KSSL Recognizer and Voice-XML" in 2006 IEEE International Conference on Fuzzy Systems. 2006, pp. 956-961.

[8] Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., & Taylor, J." Emotion recognition in human-computer interaction." *IEEE Signal Processing Magazine*. 2001: 18 (1), 32–80.

[9] Berthouze, N. Fushimi, T. Hasegawa, M. Kleinsmith, A. Takenaka, H. Berthouze, L. "Learning to recognize affective body postures" in IEEE International Symposium on Computational Intelligence for Measurement Systems and Applications, 2003. pp. 193- 198.

[10] Pook P.K, Ballard D.H.. "Teleassistance: A gestural sign language for teleoperation," *Proceedings of Workshop on Gesture at the User Interface, International Conference on Computer-Human Interaction CHI 95*, Denver, CO, USA. 1995

[11] Hauptmann, A.G. and McAvinney, P. Gestures with speech for graphic manipulation, International Journal of Man-Machine Studies, 1993. 38(2): 231-49.

[12] Wolf C. G. and Morrel-Samuels P. The use of hand-drawn gestures for text editing, International Journal of Man-Machine Studies, 1987. 27: 91-102.

[13] Rempel D, Hertzer E, Brewer R. "Computer Input with Gesture Recognition: Comfort and Pain Ratings of Hand Postures." *Human-Computer Interaction International* 2003, Vol III, Crete, Greece.

[14] Kölsch, Mathias; Beall, Andrew C.; Turk, Matthew. "An Objective Measure for Postural Comfort." *Human Factors and Ergonomics Society Annual Meeting Proceedings, Communications* , 2003. pp. 725-728(4)

[15] Wachs J. Optimal Hand Gesture Vocabulary Design Methodology for Virtual Robotic Control. PhD Dissertation, Ben Gurion University of the Negev, Israel. 2007

[16] Stern, H., Wachs, J. P. and Edan, Y. "Designing Hand Gesture Vocabularies for Natural Interaction by Combining Psycho-Physiological and Recognition Factors," *Int. J of Semantic Computing. Special Issue on Gesture in Multimodal Systems.* 2008, Accepted

[17] Wachs J., Stern H., Edan Y. "Cluster Labeling and Parameter Estimation for the Automated Setup of a Hand-Gesture Recognition System." *IEEE Transactions on Systems, Man and Cybernetics. Part A.* 2005 35:(6), pp. 932-944.

[18] Connolly D. T. "An improved annealing scheme for the QAP," *European Journal of Operational Research*, 1990, 46: 93-100.

SignWriting: Sign Languages Are Written Languages

By Valerie Sutton & Adam Frost
Center for Sutton Movement Writing (CSMW)
Sutton@SignWriting.org • Frost@SignWriting.org

Abstract

SignWriting is one of the world's scripts, ISO 15924 Sgnw, recognized by the ISO on October 10, 2006:

ISO Codes for Recognized Scripts
http://unicode.org/iso15924/iso15924-codes.html

The SignWriting Spatial Alphabet is a daily writing system for any sign language in the world. It is used by thousands of people worldwide, in some 40 countries. It can be written by hand as a daily script, or typed by computer, making it possible to publish books and web sites in written sign languages.

1. Introduction

Sign Languages are becoming written languages because of SignWriting.

Software can be accessed freely. Thousands of people create their own SignWriting dictionaries and long documents written in the handshapes, movements and facial expressions of their native sign languages, directly on the web:

SignPuddle Software Online
http://www.SignBank.org/signpuddle

The SignWriting Central Web Site provides online SignWriting Literature, Lessons in SignWriting, and an Archive of PDF documents and videos, all free for download:

SignWriting Central Web Site
http://www.SignWriting.org

First invented by Valerie Sutton, an American, in 1974, while living and working in Denmark, SignWriting was first used, in those early years, to write Danish Sign Language. Returning to her native US, Ms. Sutton founded the Center for Sutton Movement Writing in 1974, a 501c3 non-profit educational organization, now located in La Jolla, California:

Center for Sutton Movement Writing
http://www.MovementWriting.org

Lucinda O'Grady Batch, a Deaf native American Sign Language user, together with Valerie Sutton, founded the Deaf Action Committee for SignWriting in 1988, to encourage input from Deaf people who are writing their native sign languages for the first time in history. SignWriting is a successful writing system because of the influence and use by skilled signers. It is evolving naturally as more and more people write, and as Deaf children learn it in schools.

2. Purpose & Background

SignWriting is an alphabet for writing body movement, and was not developed for linguistic research, although linguistic research labs do use SignWriting to study how signers write their language.

History of SignWriting
www.signwriting.org/library/history/index.html

SignWriting was not developed with a prior knowledge of any one sign language, but was instead developed as a writing tool for writing how the body looks while signing. The system includes symbols for writing handshapes, movements, facial expressions, head movement, full-body movement, spatial relationships and punctuation. It can be used to write foreign sign languages, as well as sign languages that are well known. Signers can learn SignWriting symbols, and write any sign they know or see.

Sometimes SignWriting is compared to other linguistic notation systems for recording certain aspects of signed languages, for example the Stokoe system in 1969, and the HamNoSys system in 1988, but Sutton SignWriting was developed for a different purpose, and therefore should not be compared to linguistic notation systems. The purpose of SignWriting is to create a daily writing system for Deaf and hearing signers, who use Sign Language as their primary language, and to preserve the native sign languages around the world. Linguistic-notation-systems, such as Stokoe or HamNoSys, were not developed as a daily writing system. A comparison is on the web:

SignWriting Linguistics Forum
http://www.signwriting.org/forums/linguistics/

SignWriting is written by hand, or by computer, by signers as young as age five, or any age. It is used in Deaf Education to teach Deaf students to read and write, in Nicaragua, Germany, the USA, Canada, Belgium, Spain and other countries.

SignWriting is also used to preserve the world's sign languages, which are rich and visual languages with sophisticated grammars. Now for the first time, the beauty and deep meanings of Deaf author's poetry can be recorded on paper, and preserved for readers in the next generation.

Who Uses SignWriting?
http://www.SignWriting.org/about/who

SignWriting began by writing facial expressions as well as body movement. Facial Expressions are a part of the grammar of sign languages. Without a way to write facial expressions, one cannot write Sign Language Literature. SignWriting also records other important grammar details, such as spatial comparisons with three lanes, and the details of torso and full-body movement. So this is the first time in history that it is possible to write true Sign Language Literature. A full 140 page novel, written in the movements and facial expressions of Spanish Sign Language (Madrid, Spain), for example, would not have been possible without writing facial expressions and body movement. All 21 Chapters of the Gospel According to John have now been written in American Sign Language. Large documents written in several sign languages are being published in 2008:

SignWriting Library
http://www.SignWriting.org/library

SignWriting Literature Project
http://www.SignWriting.org/literature

American Sign Language Literature Puddle
signbank.org/signpuddle/index.html#sgn-US

3. Brief History

SignWriting stems from Sutton Movement Writing, a complete movement notation system for recording all body movement:

Sutton Movement Writing
http://www.MovementWriting.org

There are five sections to Sutton Movement Writing:

DanceWriting
http://www.DanceWriting.org

SignWriting
http://www.SignWriting.org

MimeWriting
www.MovementWriting.org/animation/mime

SportsWriting
http://www.MovementWriting.org/sports

Gesture Analysis
http://www.MovementWriting.org/science

Sutton DanceWriting was the first section of the system to be developed in the early 1970's, while Valerie Sutton was training with teachers of the Royal Danish Ballet in Copenhagen. Sutton developed a way to write dances on paper, preserving the Royal Danish Ballet's historic Bournonville ballet training system.

While Sutton was teaching the Royal Danish Ballet to read and write dance movement in 1974, a sign language researcher at the University of Copenhagen asked Sutton to record the movements of hearing and deaf people's gestures from a videotape, for a research project. Sutton became intrigued with the beautiful rich sophisticated grammar of Danish Sign Language, and became determined to develop a way to write any sign language in the world, as a daily writing system. This was the beginning of SignWriting. Upon returning to her native country, the United States, Sutton contacted Deaf people and began to work with Deaf people daily to develop a written form that is truly useful. It is because of this collaboration with Deaf native signers, that SignWriting is used today.

SignWriting is 34 years old in 2008. It was written by hand from 1974 to 1984. From 1981 to 1984, the SignWriter Newspaper, the first newspaper in history written in American Sign Language (ASL) and some articles in Danish Sign Language (DTS) was published four times a year, with articles written by Deaf native signers in the US and Denmark. It was published without any computer to help with the preparation of the 20 page newspaper. Each page was painstakingly written by hand with ink pens. It took three months to prepare one issue! Publication of the SignWriter Newspaper was halted in 1984 to focus time on developed SignWriting software to publish by computer:

History of the SignWriter Newspaper
ww.signwriting.org/library/history/hist005.html

The Deaf Action Committee for SignWriting (the DAC) was founded by Lucinda O'Grady Batch with Valerie Sutton in 1988. Deaf people, native to ASL, came together to write their language, and had, and still have, an important influence on the writing system. Today's Deaf DAC members include Lucinda O'Grady Batch, Adam Frost, Stuart Thiessen, Philippe Gallant, Kevin Clark and Darline & Dave Gunsauls:

The DAC
http://www.SignWriting.org/deaf

Videos of Deaf people discussing issues that surround SignWriting are on Google:

SignWriting Videos of Deaf Opinions
http://www.SignWriting.org/video

In 1986, Richard Gleaves developed the first software for SignWriting: a sign language processing program called SignWriter //e and //c, for the Apple //e and //c. It later became SignWriter DOS in MS-DOS and SignWriter Java. The source code for these old programs are on the web and freely available for download:

SignWriting Software Downloads
http://www.SignWriting.org/downloads

SignWriter DOS
www.SignWriting.org/forums/software/sw44

SignWriter Java
www.signwriting.org/forums/software/sw50/

SignWriter Tiger
http://www.signwriter.org

4. SignWriting Projects in 2008

2004-present
SignPuddle Software Development
http://www.signbank.org/signpuddle

SignPuddle Software is the world standard for creating dictionaries and documents in SignWriting. Documents are written in "vertical columns with lanes", directly on the web. The freedom of posting documents on the fly, in Wikipedia-like fashion, has resulted in the spreading of SignWriting all over the world.

Thousands of people from some 40 countries use SignPuddle Online on a regular basis for their students and signing communities. SignPuddle Software is the creative genius of software designer Steve Slevinski, who collaborated with Valerie Sutton, starting in 2004. The Slevinski-Sutton collaboration is ongoing.

SignPuddle 2.0 will be released August 2008, which will include Sutton's International SignWriting Alphabet (ISWA 2008), and the new SignWriting MediaWiki Plugin, which will be used on the Wikipedia site in the future. The SignWriting Image Server (SWIS) will make it possible for programmers to place SignWriting on their own web sites. The source code for the software, and the complete ISWA symbols, can be downloaded on SourceForge, under the GPL and OFL licenses.

Later in this paper, we will describe how we write signs and documents using SignPuddle Software.

2006-present
SignWriting Literature Project
http://www.SignWriting.org/literature

The SignWriting Literature Project employs skilled Deaf and hearing signers to write the world's literature in SignWriting. The initial funding for the project came from the Claire Giannini Fund. Current books being written are: The Gospel According to John in ASL (21 chapters); Sleeping Beauty, Cinderella, Snow White, Cat in the Hat by Dr. Suess, ASL Grammar Lessons in ASL, an ASL Dictionary and an ASL Encyclopedia & Wikipedia.

2002-present
SignBank Database Software Development
http://www.signbank.org/signbank.html

SignBank Software is database software for SignWriting, programmed in FileMaker Pro by Todd Duell. SignBank includes: SymbolBank, for finding and learning symbols in SignWriting, the SignBank Editor and Portal, for searching, sorting and printing dictionaries sorted by Sign-Symbol-Sequence, and DocumentMaker, a tool for publishing books in SignWriting.

1999-present
SignWriting Literacy Project
http://www.SignWriting.org/forums/teachers

The SignWriting Literacy Project began in 1999, by developing four "SignWriting Learning Levels" for teaching children. These four Learning Levels were published in ASL storybooks. The books were then donated to teachers in the Albuquerque Public Schools. Dr. Cecilia Flood, at the University of New Mexico, wrote her dissertation on the experience of teaching Deaf children using SignWriting there. Teachers at the Hodgin Elementary School documented positive experiences. Children gained pride in learning both ASL and English because now there was a way to write both languages. Other schools requested the same materials and joined the SignWriting Literacy Project. Teachers can apply online to receive free donations of SignWriting books and software.

1997-present
SignWriting Internet List
http://www.SignWriting.org/forums/swlist

The SignWriting List is an active internet list of SignWriting users with members around the world. Discussions include technical support on writing signs and symbols, how to use SignWriting software, and techniques in teaching Deaf children. Researchers also join the List to share their research. The SignWriting List Archives provide historic records of SignWriting events, and are open for searching on the web.

5. Dissertations on SignWriting

2002-2007
Dissertation, Jordan University, Ahman
by Dr. Mohammed Mahmod Abushaira
http://www.signwriting.org/arabia

Two research studies using SignWriting in Deaf education in the Middle East:
1. Research Study, Jordan 2007, Jordan University, Al-Amal School for the Deaf
2. Research Study, Saudi Arabia 2002 Alqasseem-Boraydah Deaf Children

Signing Deaf children at the Al-Amal School for the Deaf, in Jordan in 2006-2007, and in the Alqasseem-Boraydah Deaf Children Institute in Saudi Arabia in 2002, were placed in two control groups. One group learned science without a written form for their signed language, and the other group studied science with the aid of written signs in SignWriting. The group that used SignWriting had improved test scores by a wide margin in both studies. Dr. Abushaira is now Assistant Professor, Special Education, King Saud University, Riyadh, Saudi Arabia, and is writing portions of the Koran in Saudi Arabian Sign Language in SignWriting.

December, 2002
Dissertation, University of New Mexico
by Dr. Cecilia Flood

"How Do Deaf and Hard of Hearing Students Experience Learning to Write Using SignWriting, A Way To Read and Write Signs?"

Research study conducted for several years with Deaf children in three Elementary Schools in Albuquerque, New Mexico. Deaf children used the four Learning Levels of SignWriting in a series of children's books developed by Valerie Sutton and the SignWriting Literacy Project. Five classrooms of Deaf children participated, and several teachers learned SignWriting well. Dr. Flood videotaped the children learning and using SignWriting and determined that the children who used SignWriting had a sense of pride in their own language, increased self-esteem and improved communication skills in both written sign language and written spoken language. Download 300 Page Dissertation http://www.signwriting.org/archive/docs5/sw0476-CeciliaFlood-Dissertation2002.pdf

Other studies on SignWriting have been conducted in Nicaragua, Denmark, Ireland, Belgium, Switzerland, Germany, and the US. A 2008 research project is being conducted at a school for the Deaf in Quebec, Canada.

6. Writing Documents in the Visual Symbols of SignWriting in SignPuddle

Adam Frost's American Sign Language video "Why SignWriting?" is **a document written in SignWriting (see pages 6-11 below)**. It can be viewed on the web signed on video, with the written ASL transcript directly beside the video.

Go to:
Why SignWriting?

http://www.signbank.org/SignPuddle1.5/searchword.php?ui=1&sgn=5&sid=352&search=why+signwriting&type=any

When you go to the above link, you are viewing the document in SignPuddle Software. Adam wrote the transcript in the visual symbols of SignWriting directly on the web, right where you are viewing it.

SignPuddle software is free for anyone to use. With an internet connection, you can start writing signs and documents in SignWriting:

1. Download Firefox
http://www.mozilla.org

SignPuddle works best in Firefox, since it was designed for Firefox. Other browsers work too.

2. Open Firefox. Then go to:

SignPuddle Online
http://www.signbank.org/signpuddle

3. Click on the the American Flag, to access the American Sign Language Dictionary, Literature and Encyclopedia Puddles.

4. To learn how to use SignPuddle Software:

View SignPuddle Video Instruction
http://www.signbank.org/signpuddle/help/

or…

Join the SignWriting List to ask questions:

SignWriting List
http://majordomo.valenciacc.edu/mailman/listinfo/sw-l

There are many ways to learn SignWriting online:

Lessons in SignWriting
http://www.SignWrting.org/lessons

or contact us anytime!

Valerie Sutton
Sutton@SignWriting.org

Adam Frost
Frost@SignWriting.org

1. SignWriting
Read & Write Sign Languages
http://www.SignWriting.org

2. SignPuddle
Create SignWriting Documents Online
http://www.SignBank.org/signpuddle

3. SignBank
SignWriting Desktop Databases
http://www.SignBank.org

4. SignWriting List
Technical Support: Ask questions to the List
http://www.SignWriting.org/forums/swlist

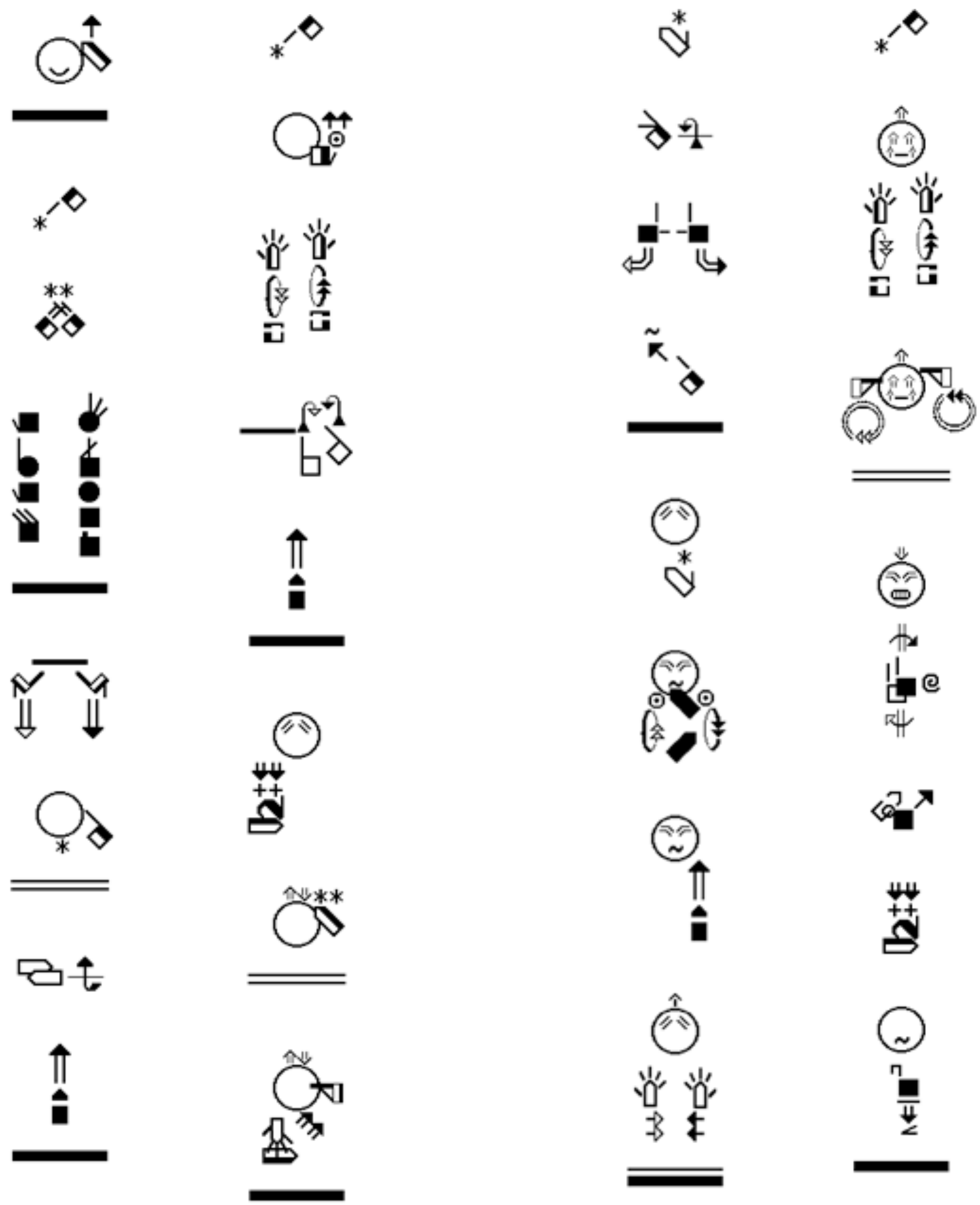
5. SignWriting Literature Project
Writing Literature in Sign Languages
http://www.SignWriting.org/literature

Deaf Action Committee For SignWriting
Center For Sutton Movement Writing
an educational nonprofit organization
P.O. Box 517, La Jolla, CA, 92038, USA
tel 858-456-0098   fax 858-456-0020
Skype Names: valeriesutton, icemandeaf
D-Link Videophone: 66.27.57.178

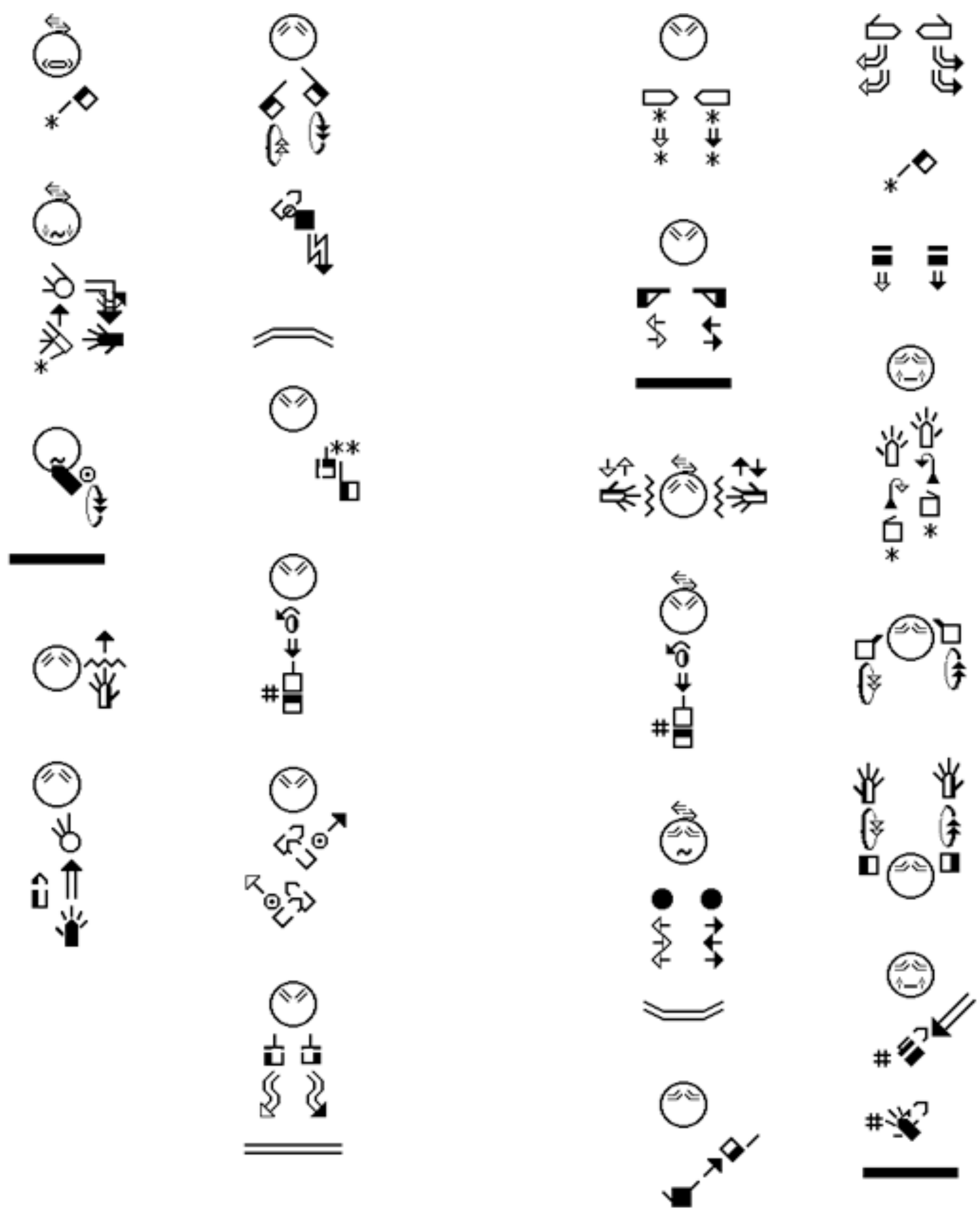
Video Transcript in SignWriting:

Hello. My name is Adam Frost. I have been deaf since birth. I use sign language everyday of my life. I have learned English.

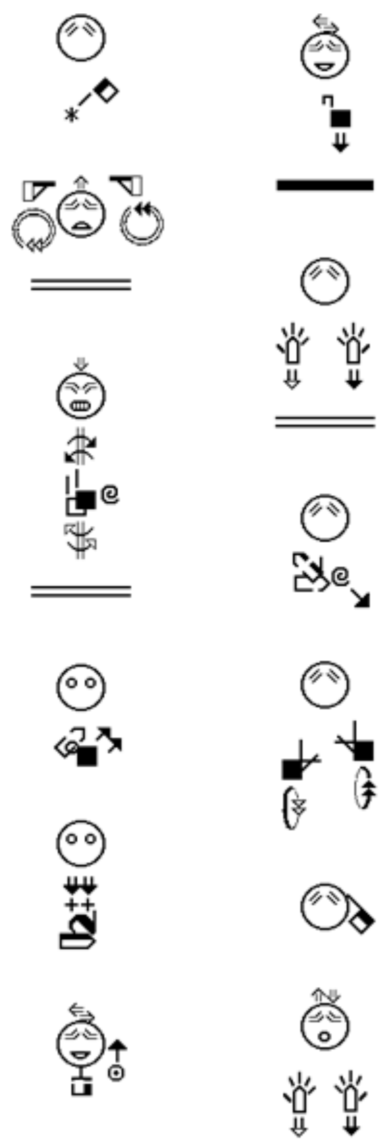
It is my second language. My greatest frustration growing up was trying to take my thoughts, which were in sign, and having to write them in English.

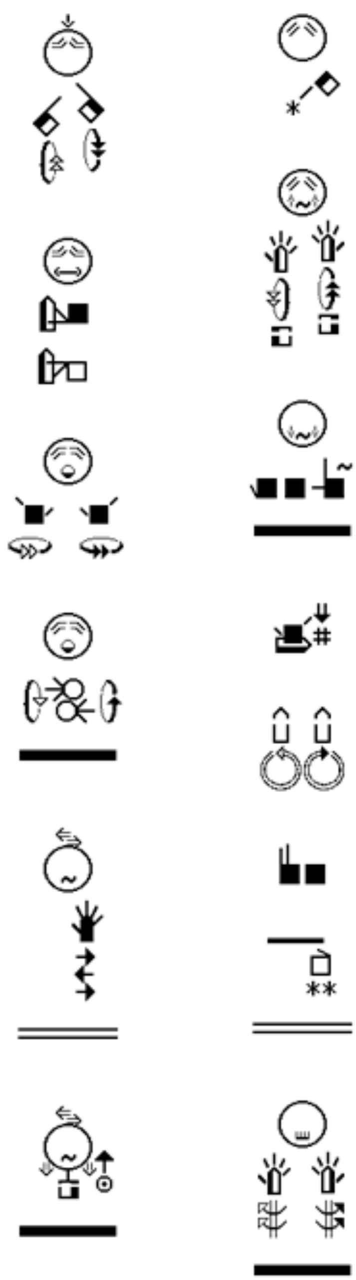


I never liked that feeling of frustration. Later when I found SignWriting, which is a system based on handshapes and … movements, not based on any sound what so ever, I found that I was able to express my ideas in sign freely without constraint on paper.

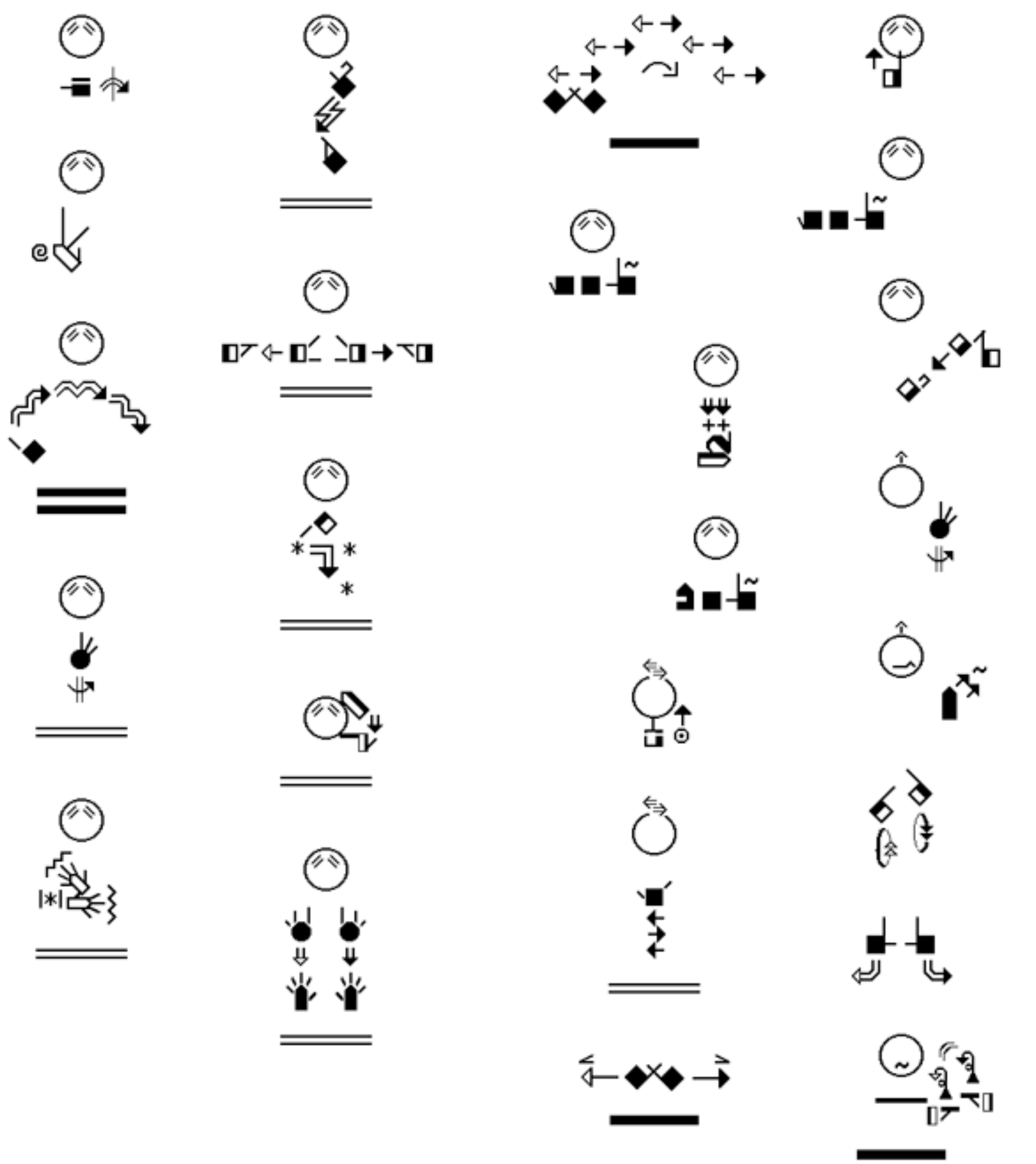


I didn't have to struggle with translating my thoughts into English just so that I could write them on paper. So, some people think that… sign language is a universal language, but that is not true. I use American Sign Language, which is only used across the United States and Canada.

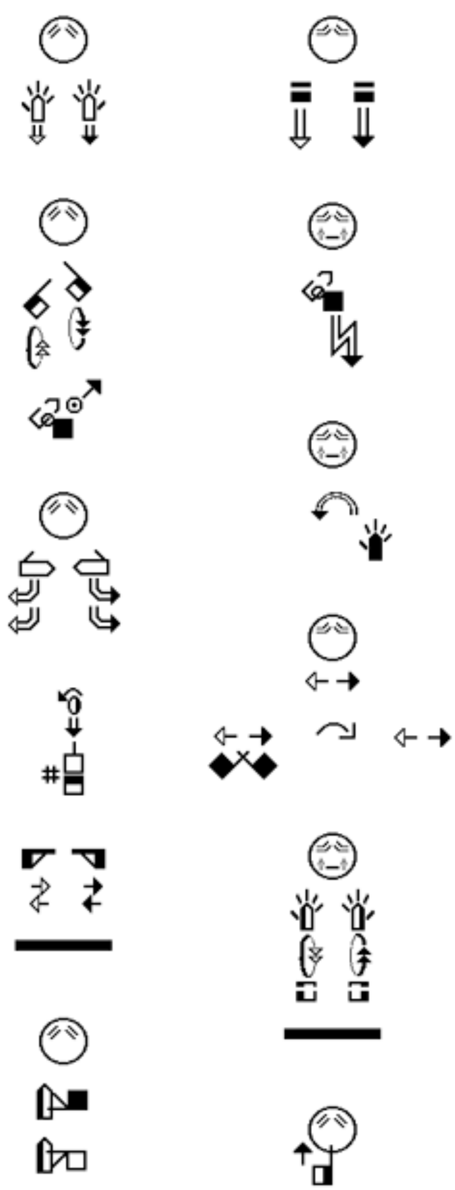




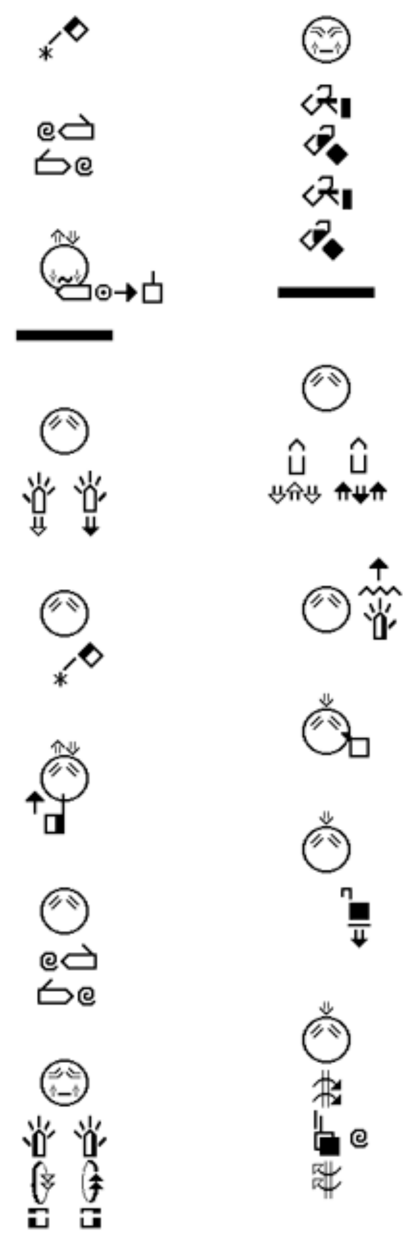
Other countries such as France, Germany, Italy, Japan, China, Korea, and Australia… all use their own sign language. In fact, American Sign Language and British Sign Language are nothing alike. In all actuality, ASL historically comes from the sign language of France.



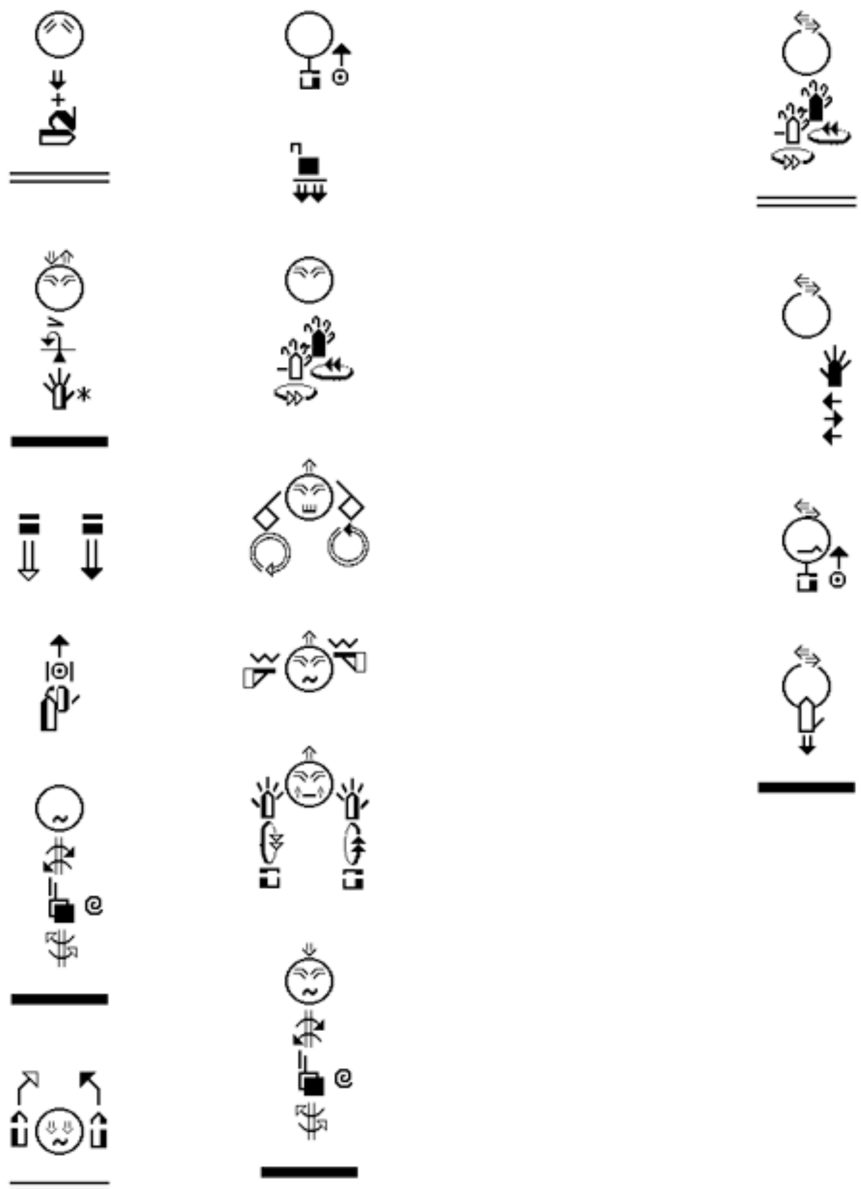
Now, because SignWriting is based on movements, that means that any of these sign languages can be written which…



I like better. Now, I am really glad that I can write directly from my native language. And if there comes a time when what I have written needs to be translated



into English, then it can easily be done so. The task of writing can be solely focused on without having to think about how to translate something because that disconnects the thought patterns.



A non-optical method for capturing the muscular-skeletal configuration of a human body part

Vincent V. Dinh, Nghia X. Tran, and Hoa V. Phan
SPAWAR Systems Center
San Diego, USA

Abstract---**Transferring human motion and likeness to a robot is a formidable task. One way to achieve this is to model the muscular-skeletal states/configurations of the human body. This information can then be transferred to a robot having those approximate states. In this presentation, we focus on just hand motions. Specifically, we are interested in the muscular-skeletal configuration of the fingers. We translate these finger data into states reflecting the muscular-skeletal configuration of finger placement. Within the range of bodily motion, these states are unique; thus, allowing them (states) to be transferred and further processed without loss of the muscular-skeletal-configuration information. To validate this property, we input the states into a rendering tool to graphically see how well the transferred states replicate the actual configuration.**

Since our technique/procedure is tractable, the whole human body may be coded in a similar manner. Conceptually, the transferred states can be used to mirror the human body via a robot---which can move with human likeness.

Keywords: **body motion kinesthetic tracking, non-optical non-triangulation data methods, motion sensors, canonical variables, group and joint angles.**

I. INTRODUCTION

In kinesthetic analysis pertaining to (computer) animation, sports training, or muscle injury diagnosis, tracking motion or body parts is the intermediate problem. This is further broken into two associated problems: forward (direct) and inverse (indirect) kinematics. The forward problem is "given the joint angles, compute the spatial configuration"; and the inverse problem "given the spatial configuration, compute the joint angles". Yamasaki et al [3] used video cameras and Polhemus sensors to track the spatial movements of the body. With similar apparatus, Badler's team [1] demonstrated joint angle extraction from kinesthetic data. Operated in a well defined confine, their kinesthetic data systems combined optics and field intensity triangulation as typically used by workers in the field.

The joint angles are often computed from under-constrained system of equations. Conventionally, the workers in the field use a minimizing algorithm as in [1]. This inherently involves iteration, rendering the procedure slow and costly.

In addition, the optical and triangulation techniques typically provide the spatial configuration data. See for example [1,2]. As mentioned, these data techniques fall in the second class of problems. Optical detection can hide motion. If the system can not see, it cannot provide mathematical kinesthetic correspondence between the representation and the subject. Manipulation can be difficult if not impossible. The limited correspondence cannot provide information on the actual 3D motion. The ideal data system should allow direct linkage (in a natural manner) to the actual motion. Either an effective method for going from the motion kinematics to joint angles is required or a new approach based on the direct problem is required.

In this paper, we suggest a new approach of getting the joint angles directly using geometrical concepts. We are interested in representations independent of spatial confinement; in particular, joint-angle representation using non-optical and spatially unconfined data. We demonstrate the link between the joint angles and the motion data. After a brief discussion of the technique, we show a typical conversion; then follow with a simple application.

II. CONCEPTS

In the above studies, when viewed as spatial configurations, the human body has many redundant degrees of freedom. Motion kinesthetic involves many variables, but many of which are interrelated. In classical mechanics, we know that these variables are functions of the canonical variables [4]. Since canonical variables reflect the degree of freedom in the system and since human motion in general is limited, there must be a finite or enumerable set of canonical variables.

To represent the state kinematics of a subject in a tractable manner, it is most effective to describe it using its natural references. These are mutually exclusive or independent references intimately tied to the familiar eigenfunctions and eigenbasis. (See any text on orthogonal function theory and abstract algebra.) For a human subject, these eigenfunctions (or eigen-states or eigen-configurations) are the joint angles. This is a concept used implicitly by workers in the kinesthetic domain. According to [1], any pose is represented by a set of joint angles. These joint angles are suggestive of the canonical variables.

In the following paragraphs, we estimate the angular displacement between two sensor readings. In the next section, we apply this to estimate the group joint angle between the index finger and the palm. We assume that the acceleration of the sensor is relatively small compared to the static gravitational field reading. If a subject is wired with many sensors, we assume uniform magnetic and gravitational fields about it. Thus, the assumption that all the sensors experience the same fields holds.

We denote the field readings by $(\vec{H}, \vec{A})$. Before we begin, we also define the directions: down, magnetic north, and east. These directions form an orthogonal set in

3D. They are estimated as follows. We first normalize the gravitational reading to form the down direction. Next we estimate the magnetic north direction with

$$\vec{H}_t = \left(\hat{A} \times \vec{H}\right) \times \hat{A} \quad (1)$$

The east direction is obtained by crossing the two previous directions. If we wanted to know the relative angle between two sensors, we compute the inner product of their orientations. For example, if we like to compare the face of one sensor to that of another, we simply take the inner product

$$\hat{n}_1 \cdot \hat{n}_2 = \cos\theta \quad (2)$$

where the faces are represented by $\hat{n}_1$ and $\hat{n}_2$; and the angular displacement between the faces is θ. In like manner, we can potentially wire a subject with a series of sensors and inquire the angular displacement of the faces of the sensors. If we place the sensors at the joints, these angles may represent the muscular-skeletal configuration of the body.

III. RESULTS

To verify the concept, we attached two motion sensors on the hand, one on the index finger above the finger nail and one on the palm as depicted in Figure 1. The idea is to translate the signals from the motion sensors to a group joint angle measure between the index finger tip and the palm. In Figure 2 below, we show the raw data from the sensors as well as the resulting group joint angle in Figure 3. The raw data are acquired over a time period of approximately $20\,s$ at $100\,ms$ intervals. The kinesthetic correspond to the index finger starting from the pointing posture then curling into a hook, then back to a pointing posture. At all times the major joint on the index finger is extended (so not to form a fist).

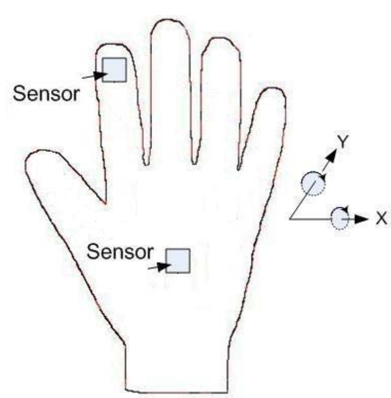


Figure 1. Sensor placement on back of hand.

Note that the correspondence between the actual motion and that of the raw data is not apparent. But after the real-time translation to the relative group joint angle between the index finger and the palm, the correspondence becomes apparent. At the start, the pointing posture has a relative group joint angle of approximately zero degree (as indicated in Figure 3). On reaching the hook posture, the relative angle approaches 180 degree. Then on returning to the pointing posture, the relative angle goes back to zero.

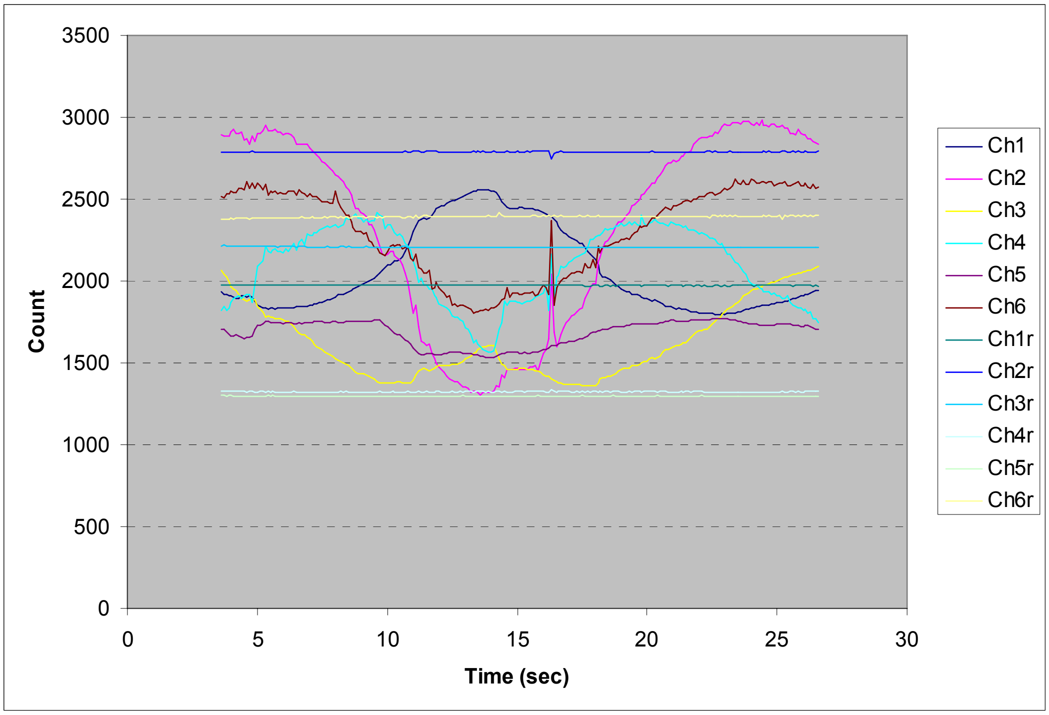


Figure 2. Raw sensor data sampled at $f_s = 10\,Hz$.

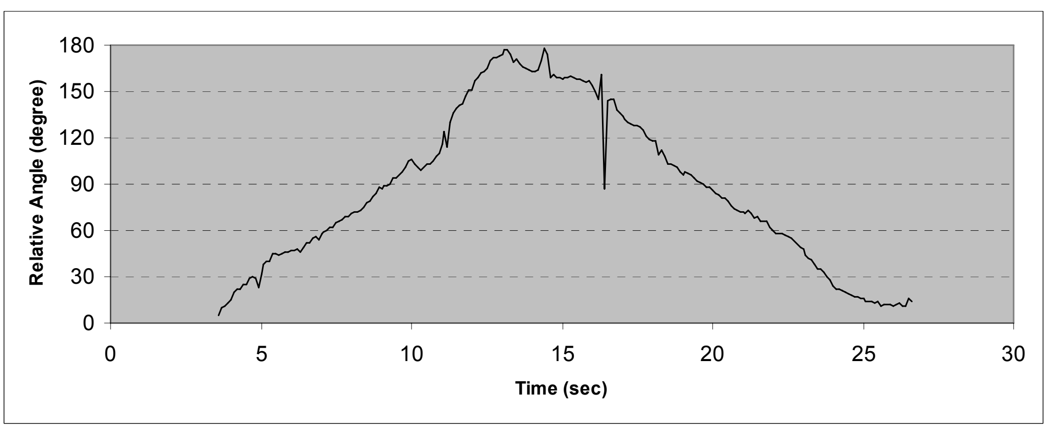


Figure 3. Computed group joint angle between index finger and palm.

Next, we demonstrate this result in an air-mouse application. The idea is emulate a standard mouse by mapping the relative joint angle of the index finger to electrical signals as produced by a standard USB mouse.

IV. CONCLUSIONS

In this presentation, we briefly described the notion of non-optical non-confined tracking. As a simple feasibility test, we computed the group joint angle between the index finger and the palm; and demonstrated its use in an air-mouse. The relative ease and usefulness of the results suggest that our techniques can be applied to many common products, such as keyboards, cell phones, remote controls, etc. This does not, however, imply that our approach is flawless. More work is needed to fine-tune our non-optical non-confined metrology techniques to the kinesthetic problem.

REFERENCES:

[1] Norman I. Badler, Cary B. Phillips, Bonnie L. Weber, Simulating Humans: Computer Graphics, Animation, and Control, 1999, Oxford University Press.

[2] Katsu Yamane and Yoshihiko Nakamura, "Natural Motion Animation Through Constraining and Deconstraining at Will", IEEE Transactions on Visulalization and Computer Graphics, Vol 9, No. 3, July-September 2003, pp352-360.

[3] Kenichi Yamasaki, Naoki Suzuki, Asaki Hattori, Akihiro Takatsu, Akihiko Uchiyama, "Skeletal and Muscular System for Quantitative Movement Analysis in 4D space", Proceedings of the 22nd Annual EMBS International Conference, July 23-28, 2000, Chicago IL.

[4] Herbert Goldstein, Classical Mechanics, 2nd Edition, 1980, Addison Wesley, Chapter 2 on normal variables.