

Enjoying Text Input with Image-Enabled IME

Toshiyuki Masui

Faculty of Environment and Information Studies
Keio University
5322 Endo, Fujisawa, Kanagawa 252-8520, Japan
masui@pitecan.com
<http://pitecan.com/>

Abstract. Tremendous amount of images are used on modern Web pages, but images are rarely used in everyday communication via e-mail, SMS, SNS, etc., although many communication systems allow the use of images in the message. We believe that images can greatly enhance the quality of communication if they are appropriately used with alphabetical texts, and we created a text input system with which users can handle images on HTML editors and word processors just like they can handle words in East-Asian languages. In this paper, we show how images are useful in everyday communication, and show how we can handle images with existing popular dictionary-based text input systems for East-Asian languages. Images are not only useful for rich communication, but they are fun to use and useful for conveying emotions.

Keywords: Text Input Systems, Image Input, Input Method Editor, IME, Dictionary-based Text Input.

1 Introduction

Tremendous number of images are used on modern Web pages for various purposes. Images are not only used for showing pictures, but they are used for showing background patterns, punctuation replacements, graphs, etc. Images are now even used in the main part of a paper like [3] for better understandings. (Fig. 1)




making direct visual references to them. For example, to ask a tour guide to explain more about a painting, we would say “*tell me more about this*” while pointing to . Giving verbal commands involving tangible objects can also be naturally accomplished by making similar visual references. For example, to instruct a mover to put a lamp on top of a nightstand, we would say “*put this over there*” while pointing to  and  respectively.

Fig. 1. Portion of the Sikuli paper

Edward Tufte is proposing the use of small graphs called Sparklines, which can be mixed in the text part of documents[2].

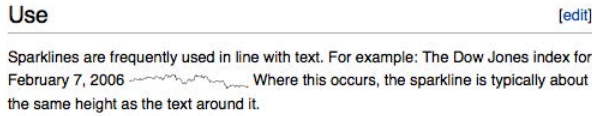


Fig. 2. Example usage of a Sparkline

Young people are getting used to using images in their e-mail messages. More than half of the Japanese Android users are exchanging HTML messages using the “Decoration Mail” feature of e-mail applications¹, with which users can use fancy images in HTML-based e-mail messages.



Fig. 3. Examples of “Decoration Mail” messages

On the other hand, elder people are still exchanging text-only messages, since composing an image-mixed text is not an easy task for them. If a user wants to put an image in his text on a word processor, he has to locate the image, copy the image to the paste buffer, and paste the image to the document. If he is using HTML, he has to save the image somewhere with an appropriate URL, and enclose it with the `` tag to show the image in the text. Various kinds of applications for “Decoration Mail” are available on mobile phones to support easy composition of fancy HTML texts, but only a small number of images are provided, and it is usually difficult or impossible to use custom images given by users.

In this paper, we introduce an image-enabled text input system, or *input method editor* (IME), with which users can input images on editors and word processors just by typing pronunciations or keywords and selecting one candidate from the candidate list generated from the input, in the same way that Japanese

¹ <http://podcast-j.net/archives/2012/04/mmd-android-ios-decome.php>

mobile phone users are entering Japanese Kanji characters. Using our IME, users can compose a text mixed with images and words in Japanese, English, Chinese and any other language, with the same input method popular on Japanese mobile phones.

2 Dictionary-Based IME

Composing Japanese and Chinese texts on PCs and mobile phones has been thought to be a formidable task, and various text input methods have been proposed and used for composing texts in those languages. Almost all the Japanese PC users are currently using variations of “Kana-Kanji conversion method” for entering Japanese texts, where a user enters the complete pronunciation of a Japanese sentence using a standard QWERTY keyboard, and the IME converts it into a corresponding Japanese text. For example, when a user wants to enter “東京駅に行きます” (I’ll go to Tokyo Station), he enters “toukyouekiniikimasu” and types the conversion key to get “東京駅に行きます”.

Typing characters like “toukyouekiniikimasu” without an error is not very difficult using a standard QWERTY keyboard, but it is not easy when using a small keyboard on a mobile phone. So, Japanese mobile phone users are using simpler dictionary-based predictive text input systems with which only a small number of keystrokes are required for entering words. When a mobile phone user types “touk”, words like “東京”(Tokyo) and “東京駅”(Tokyo Station) are listed as candidate words, and the user can select one from the list. With this method, users have to select a word in a sentence one by one, but using a good prediction algorithm, the number of keystrokes required for the user is reduced dramatically.

Fig. 4 shows how a sentence is composed using the POBox text input system[1] on an Japanese mobile phone introduced in 2003. When a user types a character “お”(o), candidate words are listed at the bottom of the display so that the user can select one of them if he finds the word he wanted to enter. After selecting a word, next input word is predicted from the selected word and listed as new candidates.



Fig. 4. POBox on a 2003 mobile phone

Dictionary-based IMEs are widely used for composing East-Asian languages, but the same technique is useful for European and other languages, and even for programming languages. The Emacs editor has the “abbreviation” feature, where users can explicitly define short abbreviation strings for long words. For example, if the user defines an abbreviation “ab” for “abbreviation”, he can type the Ctrl-X key and single quotation(’) key after typing “ab” to generate “abbreviation”. Using an abbreviation dictionary, he can type “ma” to get “Massachusetts”, type “mo” to get “Missouri”, etc. Emacs also has the “dynamic abbreviation” feature, where users can enter a long word just by typing the Meta-/ key after typing the first several characters of a word which appear somewhere in the text. For example, when a user edits this text and type “ab” and type Meta-/, “ab” will be expanded to “abbreviation”, because this text contains the word “abbreviation” which begins with “ab”. In this case, the text under composition is used as the dictionary for expanding a prefix of a long word.

The idea behind these features are almost the same as Japanese IMEs on mobile phones. The difference is that IMEs for Asian languages are heavily used for entering various texts, while static and dynamic abbreviation feature is invoked by the user only once in a while.

We have created an IME which supports entering images in the same interface as entering words in Japanese and other languages. When a user enters the first part of the pronunciation of a word or a image, candidate words and images are displayed in the IME’s candidate list, and the user can select one from the list and paste it into the text.

3 Image-Enabled IME

3.1 Using UTF Image Characters

Punctuations, exclamation mark, question mark, and other symbols are used in English texts, and face marks (e.g. “:-)”) are used everywhere these days. In addition, various symbolic characters are defined in UTF, and we can use UTF characters like ☺, ☹, ☺, ☂ for fun. If we define a pronunciation to each

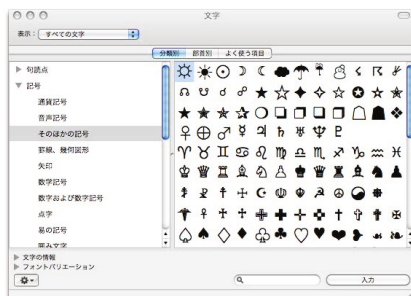


Fig. 5. UTF symbol characters available on Mac

character, we can use it in dictionary-based IMEs. For example, if we define a pronunciation “rain” to ☂, we can enter the character just by typing “rain”, as shown in Fig. 6.

Using an IME with appropriate dictionary, we can enter a UTF character like ☂ just by typing “rain”.

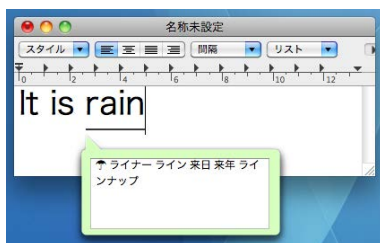


Fig. 6. Typing “rain” to get “☂”

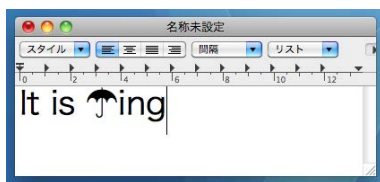


Fig. 7. Selecting “☂” by typing the space key, and continue entering texts

3.2 Entering Images

Fig. 8 shows how we can enter an image of a fish on a word processor (TextEdit on Mac). When we type “sakana” (fish) in our IME, we can see the Kanji character “魚” (fish) and other fish images displayed as candidate words, since the pronunciation “sakana” is defined for the fish images.

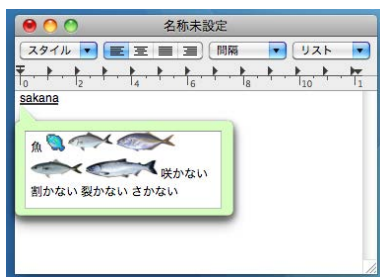


Fig. 8. Entering “sakana” to get a list of fish images

When we type the space key, the first candidate (“魚”) is selected and put into the text editing area.



Fig. 9. Selecting “魚” by typing the space key

We can select the candidate by typing the space key and the backspace key. When we type the space key n times, we can select the n -th candidate and show it in the text area. When we type the return key, the selection is fixed and the candidate list disappears.

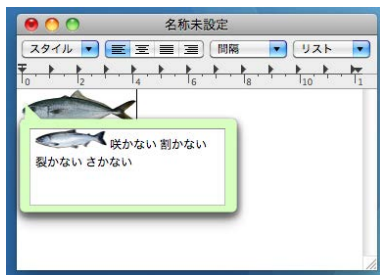


Fig. 10. Selecting tuna by typing the space key several times

In this IME, we can select an image and put it into word processors just like we enter Japanese words. Composing a text with images is as easy as composing a Japanese text.

4 Examples

In this section, we show various examples of using our IME for fun and for practical purposes.

4.1 Using Enhanced Punctuations

Question mark(“?”), exclamation mark(“!”), and other punctuation symbols have long been used for adding extra meanings and emotions to sentences. Using our IME, we can use various images for expressing feelings.

If we define a pronunciation “surprised” to images of surprised faces and put them into the IME dictionary, we can get a list of surprised faces by typing “surp” (Fig. 11), and select one of the surprised faces and paste it in the text (Fig. 12), just like we can enter Japanese words into the text area.

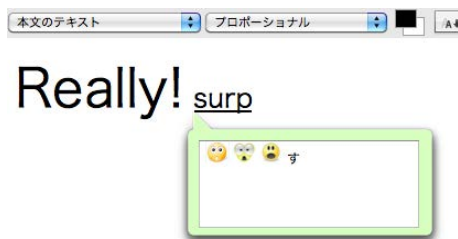


Fig. 11. Showing surprised faces by typing “surp”



Fig. 12. Selecting one of the surprised faces and pasting it into the text

4.2 Intuitive Expression

Sometimes images are easier to understand than text symbols. When you want to have a meeting between 14:00 and 16:00, you can write a message like “Let’s have a meeting at 14:00 today”. However, if the recipient didn’t read the message carefully, he might come to the meeting at 4:00pm instead of 14:00. If we use clock symbols instead of numbers like Fig. 13, nobody can make a mistake of this sort.



Fig. 13. Using clock images for specifying time

4.3 Using Faces Instead of Using Names

Instead of saying “Lena is coming today”, we can use her face, if “Lena” and her image are defined in the dictionary. This is another example where using an image is more intuitive than using a text.

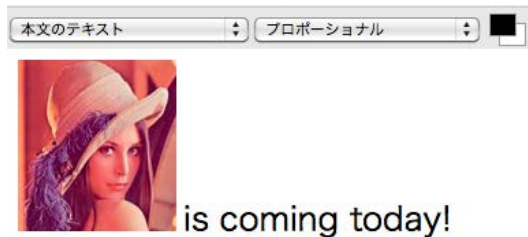


Fig. 14. Lena is coming today

4.4 Using Internet Search

Basically, all the images and words should be stored in a local dictionary, but we can also use a service like Google Image Search² for finding images of celebrities and famous places. Since it is almost sure that we can find President Obama’s face from the Internet, we don’t have to register it in the dictionary beforehand. (Fig. 15)

4.5 Dynamic Image Creation

We can also use images dynamically generated from the parameters given by the user. In the example shown in Fig. 16, the user is trying to enter an image which represents the RGB parameter. When the user enters “0000ff#”, a blue rectangle image corresponding to the parameter is dynamically generated and listed as a candidate. (Fig. 16)

² <http://www.google.com/imghp>



Fig. 15. Entering “obama!” to get images of President Obama

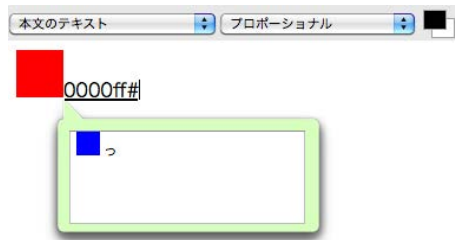


Fig. 16. Generating an image from RGB value

In the same way, it is possible to generate a Sparkline image or an analog clock image from the parameters given by the user.

4.6 Composing Attractive E-Mail Messages

Using our IME, we can easily select beautiful images and paste them to e-mail message (Fig. 17). A message with beautiful images are much more attractive than a text-only message. People don't use images in e-mail communications just because they cannot enter images as easily as entering texts. Just like Web pages became popular to the public after the introduction of the Mosaic browser which could display images on Web pages, we expect that people use more images in their everyday communication if input systems like ours become popular.

5 Implementation

5.1 Handling Images in IME

Our IME is implemented in MacRuby, using the IMKit text input library on MacOS. Since IMKit does not support image handling, image data is copied to

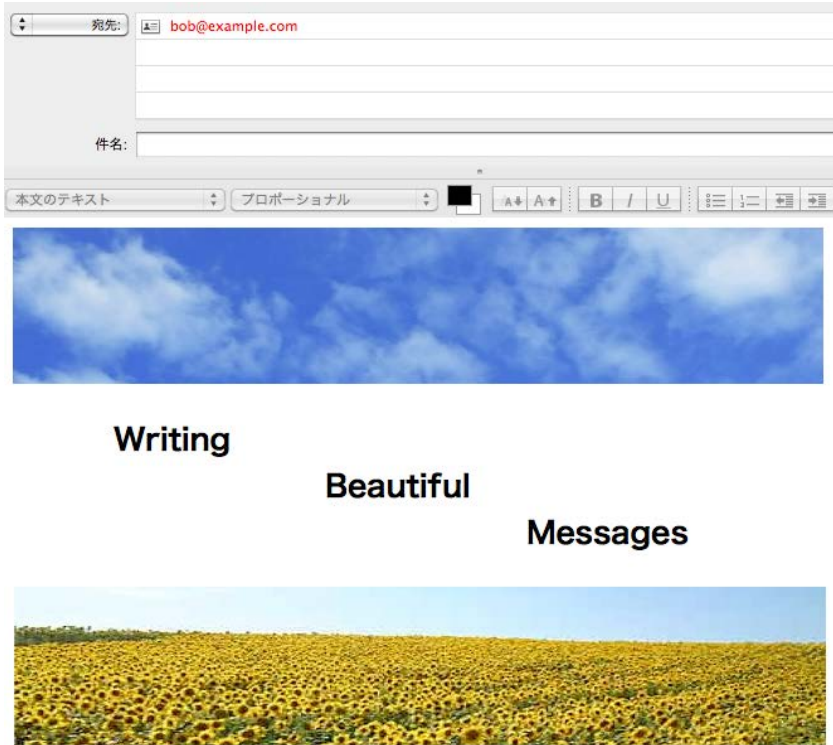


Fig. 17. Composing an e-mail message with beautiful images

the paste buffer and then pasted to editors and word processors every time it is selected by the user.

5.2 Registering Images in the Dictionary

We are using the data on Gyazo image upload service³ for the images handled in the IME. All the image data on Gyazo have unique MD5 IDs calculated from the image data, and the images used in our IME are cached in an image folder. If a user wants to enter an image from its pronunciation, he should register the pair of the pronunciation and the ID in the dictionary.

We also provide a way to upload a clipped image on the desktop and register it with the pronunciation. If we are browsing the ACE2012 Web page and we want to use the image in the IME, we can invoke the clipping/registering application (Gyazo), specify the clipping area, and enter the pronunciation for the image (Fig. 18).

³ <http://Gyazo.com/>



Fig. 18. Registering the ACE2012 icon with pronunciation “ace”

After the image is registered in the dictionary, we can type “ace” to find the image and paste it to the application.

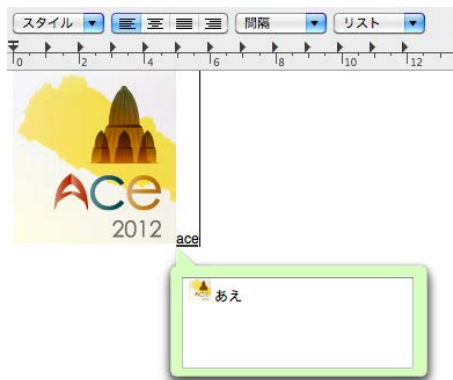


Fig. 19. Using the icon of ACE2012 in TextEdit

6 Conclusions

Receiving a message with beautiful images is pleasant, but composing a message with images has been a pain. Composing a message in Japanese and Chinese used to be a big pain more than 10 years ago, but it was alleviated after the introduction of dictionary-based IME, and now everyone is exchanging Japanese and Chinese messages between mobile phones without pain. Using a dictionary-based IME for entering images, we hope we can enjoy exchanging messages full of images.

The Japanese word “楽” has two meanings: “easy” and “to enjoy”. People have been trying to develop easy-to-use IMEs for many years, but we think we are now able to create IMEs with which we can enjoy text and image input tasks.

References

1. Masui, T.: An efficient text input method for pen-based computers. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 1998, pp. 328–335. ACM Press/Addison-Wesley Publishing Co, New York (1998), <http://dx.doi.org/10.1145/274644.274690>
2. Tufte, E.: Beautiful Evidence. Graphics Press (2006)
3. Yeh, T., Chang, T.H., Miller, R.C.: Sikuli: using gui screenshots for search and automation. In: Proceedings of the 22nd Annual ACM Symposium on User Interface Software and Technology, UIST 2009, pp. 183–192. ACM, New York (2009), <http://doi.acm.org/10.1145/1622176.1622213>