

Universal Multiple-Octet Coded Character Set

UCS

ISO/IEC JTC1/SC2/WG2 IRG N [1168](#)

Date: 2005-11-29

| | |
|---------------|-------------------------------------|
| Doc. Type: | Member body contribution |
| Title: | Principles of Old Hanzi Selection |
| Source: | China |
| Status: | Input to IRG |
| Action: | |
| Distribution: | IRG Members and Ideographic Experts |
| Reference: | |
| No. of pages: | |
| Medium: | Electronic |

China experts discussed the selection and the procedure of Old Hanzi with experts of Taiwan in the meeting held on Oct. 26-28, 2005 in Taiwan. Both parties agreed a consensus should be made before the processing of Old Hanzi, and a further cooperation is expected in the future work.

Chinese delegation has the following proposition:

1. The principle of Old Hanzi Selection:

The Old Hanzi should have a complete coverage and an exact map of the attributes of sources. Complete coverage is to guarantee the multiple use of Old Hanzi (research, publishing, virtual library).

To be complete, the glyphs to be selected should include all existing individual glyphs of all public materials.

Characters to be selected should include:

- A. different characters
- B. character with different construction
- C. character with different aesthetical style
- D. all components, their variants, varying writing forms
- E. signs of function (repetition symbol, sentence symbol etc.)

Attributes of sources should include:

- A. sources of characters to be selected.
- B. sources of glyphs covered by selected characters.

2. Procedures:

To facilitate the processing of Old Hanzi, a database should be built. It should have the following platforms:

A. Contextual information

The text in which the Old Hanzi appears, the attribute of region, time, scholarly research work, and other related information. (see fig. 1)

B. Glyph information

The original image, decomposing component, representative glyph, corresponding modern character (if any), coverage index of the selected glyph. (see fig. 2)



Figure 1



Figure 2