Associazione Investire in Cultura

FONDAZIONE ROMA MEDITERRANEO

4th International Congress on

"Science and Technology for the Safeguard of Cultural Heritage in the Mediterranean Basin"

Cairo (Egypt)
6th – 8th December 2009

B3/6 - EXTRACTING ERASED TEXT FROM PALIMPSESTS BY USING VISIBLE LIGHT*

Emanuele Salerno, Anna Tonazzini
Consiglio Nazionale delle Ricerche
Istituto di Scienza e Tecnologie dell'Informazione, Pisa, Italy
{emanuele.salerno,anna.tonazzini}@isti.cnr.it

Recently, high resolution specialized imaging has been giving a substantial help in reading palimpsested manuscripts without using invasive practices, often harmful for historically important ancient documents. In particular, faint traces of the erased texts can be highlighted by ultraviolet illumination (Wood lamps). By image processing techniques, the fluorescence images thus obtained can complement the reflection or transmission images under visible-light illumination to further help the scholar in distinguishing the erased text from the overwriting. Fluorescence images are not always available to be processed. In some cases, however, we have observed that a suitable processing of the standard color channels can help in distinguishing the two texts even though no additional channel is available.

In general, different inks show different spectral signatures, thus the two texts to be distinguished may tend to fade over different wavelength ranges. This is how multispectral or hyperspectral imaging using both reflection and fluorescence, and covering the visible plus some of the infrared and ultraviolet bands, can enable us to separate the texts. Indeed, by inspecting the channel images, one can find some wavelength range where one of the two texts disappears, whereas the other has still a good contrast. By exploiting this feature, a pseudocolor image can be built, where one of the two texts is perceptually well separated from the other (e.g. it might appear in a totally different color). Any human observer would now be capable to decorrelate the two text patterns with no further mathematical image processing.

If a standard color image is only available, both patterns are normally visible in all the channels, and the strategy described above cannot be applied so straightforwardly. However, spectral diversity can still be exploited to improve the readability of the underwritten text through a pseudocolor image that can be very similar to the one obtained by using a fluorescence channel. The key to the procedure we are proposing is the conjecture that the individual text patterns appearing in the document are statistically independent (or, at least, uncorrelated), whereas the color channels, where all the features are mixed, are strictly correlated to one another. If, as we hypothesize, the color appearance of the document comes from a linear (though unknown) mixture of all the individual patterns, the output images obtained by decorrelating numerically the color channel information are copies of the individual patterns mixed in the original. Diverse mathematical

* This work was supported by the Calabria Region PIA 2008 project no. 1220000119 AMMIRA (Multispectral acquisition, enhancement, indexing and retrieval of artworks)

approaches can be adopted to reach this goal. From our experiments, even when the linearity and decorrelation hypotheses are not verified, it can happen that the underwritten text is totally separated from all the other patterns. This would solve the problem under an image processing point of view. In some cases, however, the text to be read can still appear to be superimposed to other features. Also, scholarship needs could suggest to display all the features together. In these cases, some of the output channels can be used to build a pseudocolor image where everything is visible, but the interesting features are highlighted and well distinguishable from the interfering patterns. Our experimental results show that, by following this strategy, the readability of palimpsests can be highly improved by digital image processing even when specialized image capture facilities are not available.