Information separation in art investigation; a survey

Jan Blažek, Barbara Zitová.

Dept. of Image Processing, Institute of Theory of Information and Automation, The Czech Academy of Sciences, Czech Republic

***Abstract*— The goal of artwork analyzes is often to detect of pentimenti, retouches, overpaintings, or varnishes in order to understand a painting structure. A common model of a painting used for interpretation of an artwork multimodal dataset is based on its multilayer characteristics. Another possibility how to address an artwork structure is to study an *information gain* of a particular modality. We have developed a new approach [2] for the *information gain* extraction and demonstrated its applicability. We present a comparison of four methods for the information separation [4, 1, 3, 2] applied on a multimodal dataset. Their ability to uncover concealed features of paintings will be presented together with their requirements and limitations. The separation limits will be shown using a concept of the intensity correspondence matrix (ICM), which can well describe the correlation and the mutual information. ICM also gives evidence of possibility to achieve an effective signal separation.**

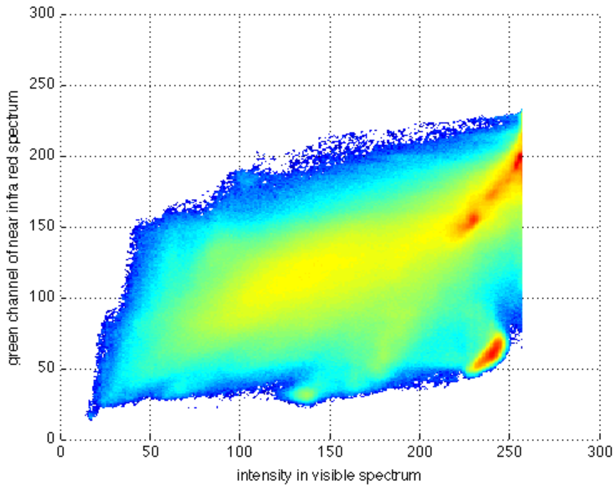

Figure 1: A 2D histogram of the intensities correspondence. In the area of VIS intensity level 230 there are two peaks. The low intensity in NIR corresponds with underdrawings while the high intensity come from areas without underdrawings. Such dataset is separable.

1 Introduction

The general problem of multimodal datasets (if we neglect the major problem of their registration) is their very high cross-modal correlation caused by the principle of reflection or transmission measurements. The information content of a painted surface affects all radiation passing through; reflectance in the visible part of the spectra (VIS) as well as penetrating modalities (terahertz (THz), X-Ray (RTG), near-infrared (NIR)). The contrast of deeper layers is significantly lowered and often falls down to the level of noise. Thanks to this, the modal images are often hardly readable. Moreover, they can go to be completely useless for art investigation (there are exceptions from this concept e.g.[5]).

The identification of features of the covered painted layers is relevant task for image processing. For this purpose the methods for a signal separation come to the scene. Making an assumption, that the VIS modality is less penetrating than the THz, the X-ray or the NIR, respectively, we can assimilate the idea of painted layers. While the VIS modal image is affected just by the surface "layers visible in VIS modality" the more penetrating modalities can be affected also by "some deeper layers" too. For simplicity, we split the painting into two parts: *the surface* or *top layer* which represents the layers affecting VIS reflectogram and *the layer underneath* which contains everything else.

In our survey we would like to compare four methods [4, 1, 3, 2] used for the multimodal dataset separation in order to visualize concealed features hidden in the images obtained in penetrating modalities.

2 ICM and its patterns

We offer a new perspective to signal separation by an *intensity correspondence matrix* (ICM), which can be used as the common denominator for all mentioned studies. Moreover, the patterns for ICM describe the problems and the limits of the information separation of a multimodal dataset.

The ICM is a matrix, which contains the frequency of correspondent pixel(s) intensities of two different modalities. As a matrix *hyper-column* we name a VIS vector while a *hyper-row* denotes a vector in a target modality. The number of hyper-rows and hyper-columns corresponds to the intensity levels recognized in each modality, while the dimensionality of the hyper-row and hyper-column is given by the pixel vector length and the pixel neighborhood size taken into account. E.g. for two modalities with intensity levels $l \in L = \{0, 1, 2, ..., 255\}$ the ICM is 2D histogram with $\|L\| = 256$ bins in both dimensions (see Figure (1)) while for approach in [3] we have 6D histogram with $(n \times d \times \|L\|)^2$ bins, where n is the number of pixels in the patch and d the length of per pixel intensity vector.

The visualization (if possible) of the (low) dimensional ICM can give us a notion of the potential separability effectiveness.

In the ICM, it is easier to recognize more probable correspondences of modal vectors from the less probable ones in the context of modality. In general, the dataset defines a **mapping** between VIS and the second modality. But an algorithm for separation is just a **function** (see Figure (2)). The discrepancy between the mapping and the function causes that all the corresponding intensities in the target modality (for one *hyper-column*) are reduced by the separation function to just one output vector. If we ask for the effectiveness of mapping to function reduction we can recognize peripheral but relevant patterns of the ICM:

- The good case - the highest peaks in ICM hyper-columns correspond to the *top layer* effect in both modalities

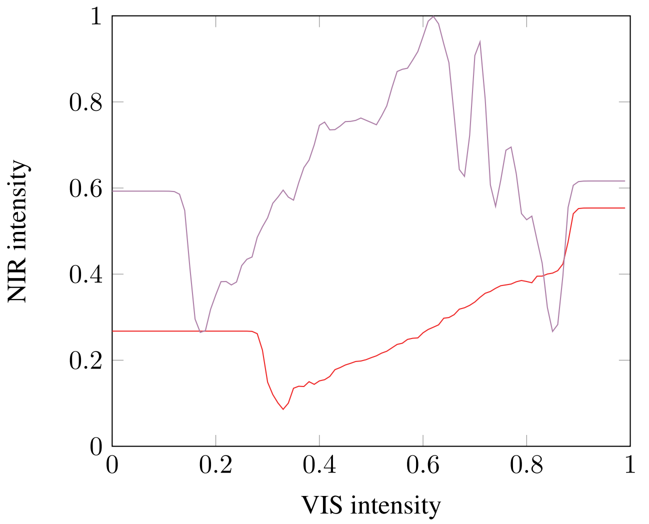

Figure 2: Two successfully trained transfer functions from the intensity in the VIS to intensity in the NIR. The purple line demonstrates how non-trivial such function can be. On the contrary, the red one is an example of the identity like function. Both lines were trained on the dataset presented in [2], the purple one on the dataset in the Figure (10), the red line on the dataset in Figure (9). The VIS input were reduced to the one dimensional intensity gray values.

- The bad case - multiple peaks - several relevant peaks per hyper-column with similar frequencies
- The bad case - smooth distribution - uniform distribution of values per hyper-column

Our results [2] highlight the problem of separability based on the number of materials and their mixtures in the dataset as well as the importance of correlation of both input modalities.

3 Generalization

We connect the presented approaches with the model of ICM, because the ICM and our approach [2] can distinguish cases when the information obtained from the NIR or the X-ray can be effectively separated and when this is impossible.

Firstly we reduce of the problem of the estimation of the separation function to the problem of searching corresponding vector in NIR or X-ray for a vector in VIS. This means an assignment of the hyper-row to a hyper-column in ICM. An assignment of these vectors to all hyper-columns defines the approximation function converting the *top layer* information in the VIS modality to the *top layer* information in the second modality.

In the Gooch's and Tumblin's paper [4] authors estimate the target function for each mean-shift based segment. For the whole segment just one X-ray vector is defined, which is computed as the mean value of the segment in the X-ray. This corresponds to the assignment of an hyper-column to the mean value of relevant hyper-rows. For the application of the method there must be just one significant peak per each hyper-column, otherwise, the mean value which is probably irrelevant will be taken as the X-ray representation.

The method in our paper [2] for a hyper-column extracts the hyper-row which minimizes the square error of approximated and real vectors. In an ideal case this is the most probable hyper-row. Our experiments demonstrated that this peak must be significantly higher than any other hyper-row frequency in the same hyper-column.

In the paper of Anitha et al. [1] the mutual entropy of an approximated *top layer* is minimized and the entropy of XRF *information gain* is maximized. This approach, in most cases, takes the highest peak in the hyper-column as the XRF representation. Other peaks in the hyper-column and low frequency values around them are put into the *under-paintings* category. Other selections of the hyper-column representative are also possible, mostly due to the applied regularization, but they need custom explanation and their stability is lower.

The unsolvable problem for our [2] approach is an existence of more than two peaks per hyper-column. Because in such case, false positives will be produced in the separated signal. A robustness of Anitha's et al. [1] approach is here improved by the wavelet decomposition which causes that a pixel neighborhood also affects output C_u and C_s intensities. But the pixel neighborhood itself does not influence the estimation of an approximation function.

The last approach of Deligiannis et al. [3] includes the influence of a pixel neighborhood into the approximation estimation. The authors use the ICM not only per one pixel intensity vector but they include into one hyper-column/row pixel neighborhood. With the raw pixel intensities this will cause a curse of dimensionality which in the extreme case can produce an unstable algorithm. The authors solve this using the coupled dictionaries with limited number of words and limited linear combination of those words. A notion what is happening here is more difficult, because more dimensions with different meaning are included into our ICM concept. In the optimal case the more dimensions and the different pixel context moves the peaks in hyper-space, defined by hyper-column dimensions, close to each other and a more relevant representative vector could be selected. However, the evaluation of this hypothesis is out of the scope of our paper. In general, stability of this approach can be negatively affected more than all previous methods by decreasing common information in analyzed modalities due to the minimization in higher dimensional space. As we presented in our study [2] the maximum of the second modality *information gain* is around 10%. This condition is for X-ray and XRF modalities hard to meet. In our case, these 10% limit also includes the noise from both modalities.

4 Conclusion

The separation of the modality information to the *top layer* and the *information gain* is in all referred studies done, in principle, by an approximation function. This function estimates the general mapping between modalities which we analyze by the ICM. The computed function in the ICM is realized by hyper-column $\rightarrow$ hyper-row pairs. We have pointed out that there exist patterns in the ICM which cannot be separated by any function, but the ICM in such cases can be constructed in different way. In [3] the ICM is constructed with respect to the pixel neighborhood while the sparsity of such ICM is reduced by dictionary code-words and patch representation rules.

For further research an identification of the bad patterns of ICM as well as an identification of methods limits is crucial. We recommend to test separation methods on phantoms, where the statistical ground truth would be known and non-trivial.

References

[1] Anila Anitha et al. "Restoration of X-ray fluorescence images of hidden paintings". In: *Signal Processing* 93.3 (2013). Image Processing for Digital Art Work, pp. 592–604. ISSN: 0165-1684. DOI: 10.1016/j.sigpro.2012.09.027. URL: http://www.sciencedirect.com/science/article/pii/S0165168412003659.

[2] Jan Blažek et al. "Improvement of the visibility of concealed features in artwork NIR reflectograms by information separation". In: *Digital Signal Processing* 60 (Jan. 2017), pp. 140–151. DOI: 10.1016/j.dsp.2016.09.007.

[3] N. Deligiannis et al. "Multi-Modal Dictionary Learning for Image Separation With Application In Art Investigation". In: *IEEE Transactions on Image Processing* PP.99 (2016), pp. 1–12. ISSN: 1057-7149. DOI: 10.1109/TIP.2016.2623484.

[4] Jack Gooch, Amy and Tumblin. "Visualizing pentimenti: revealing the hidden history of paintings". In: *Journal of Mathematics and the Arts* 00.00 (2007), pp. 1–11. ISSN: 17513472. DOI: 10.1080/1751347YYxxxxxxxx.

[5] David Thurrowgood et al. "A Hidden Portrait by Edgar Degas." In: *Scientific Reports* 6 (2016), p. 29594. DOI: 10.1038/srep29594. eprint: arXiv:1011.1669v3. URL: http://www.nature.com/srep/2016/160804/srep29594/full/srep29594.html.