Supplementary Material

Rediscovering the lost colors: film color restoration by hyperspectral imaging and Cluster-Based Spectral Correction Algorithm (CBSCA)

Emilio Catelli,[a] Lingxi Liu,[b] Jacopo Fadanni,[c] Jost Stergar,[d,e] Matija Milanic,[d,e] Giorgia Sciutto,[a] Francesco Zerbetto,*[f] Silvia Prati,*[a]

a. *Department of Chemistry "G. Ciamician", University of Bologna - Ravenna Campus, Via G. Guaccimanni 42, Ravenna, Italy. E-mail: s.prati@unibo.it*

b. *Department of Art History, Musicology and Theatre Studies, Ghent University, Sint-Pietersnieuwstraat 41, 9000 Ghent, Belgium.*

c. *Department of Physics and Astronomy, University of Padova, Via F. Marzolo 8, 35131, Padova, Italy.*

d. *Jozef Stefan Institute, Jamova cesta 39, SI-1000, Ljubljana, Slovenia.*

e. *Faculty of Mathematics and Physics, University of Ljubliana, Jadranska ulica 19, SI-1000, Ljubljana, Slovenia.*

f. *Department of Chemistry "G. Ciamician" University of Bologna, Via P. Gobetti 85, 40129, Bologna, Italy. E-mail: francesco.zerbetto@unibo.it*

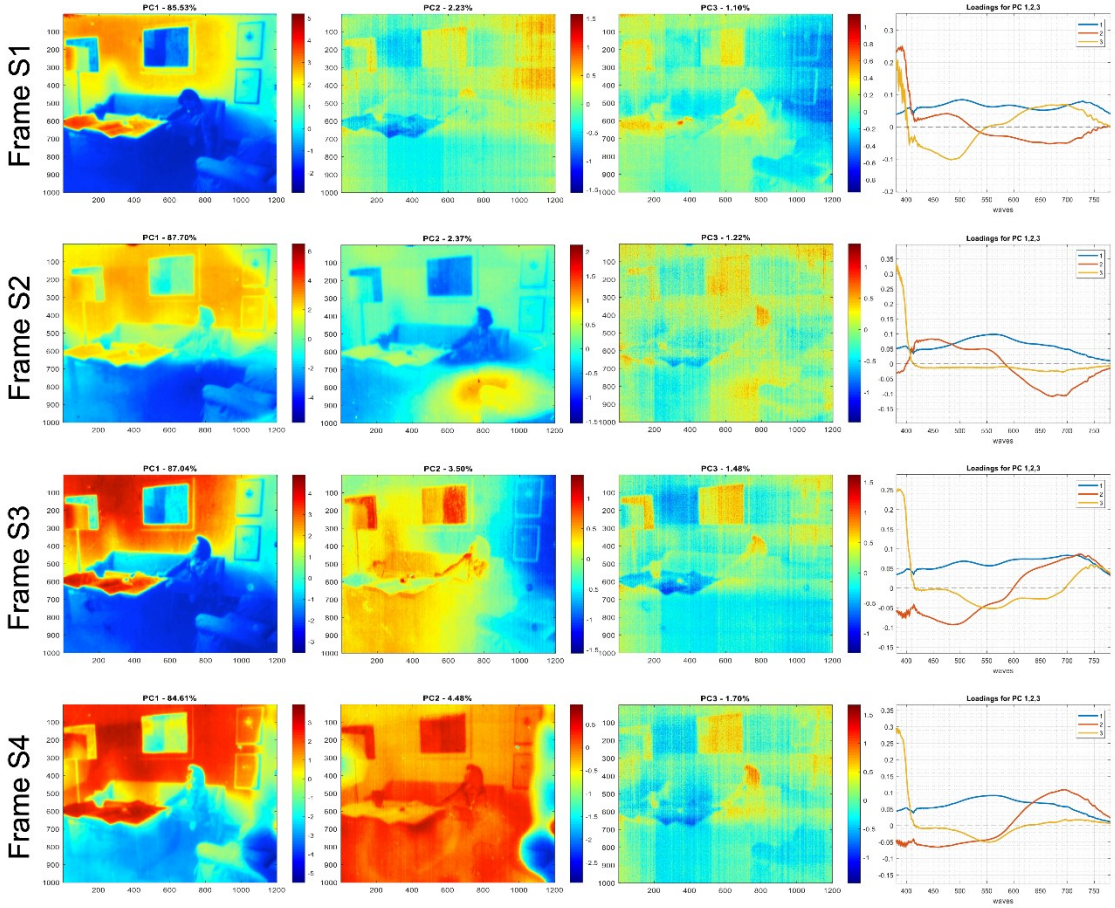

Fig. SM1 Principal Component score images and loading profiles for the S1-S4 frames. The first three components account for more than 88% of the total variance, indicating that the dominant variability of the hyperspectral data lies in a three-dimensional subspace. The data were mean-centered prior to PCA. The data were processed using hypertools (https://www.hypertools.org/)

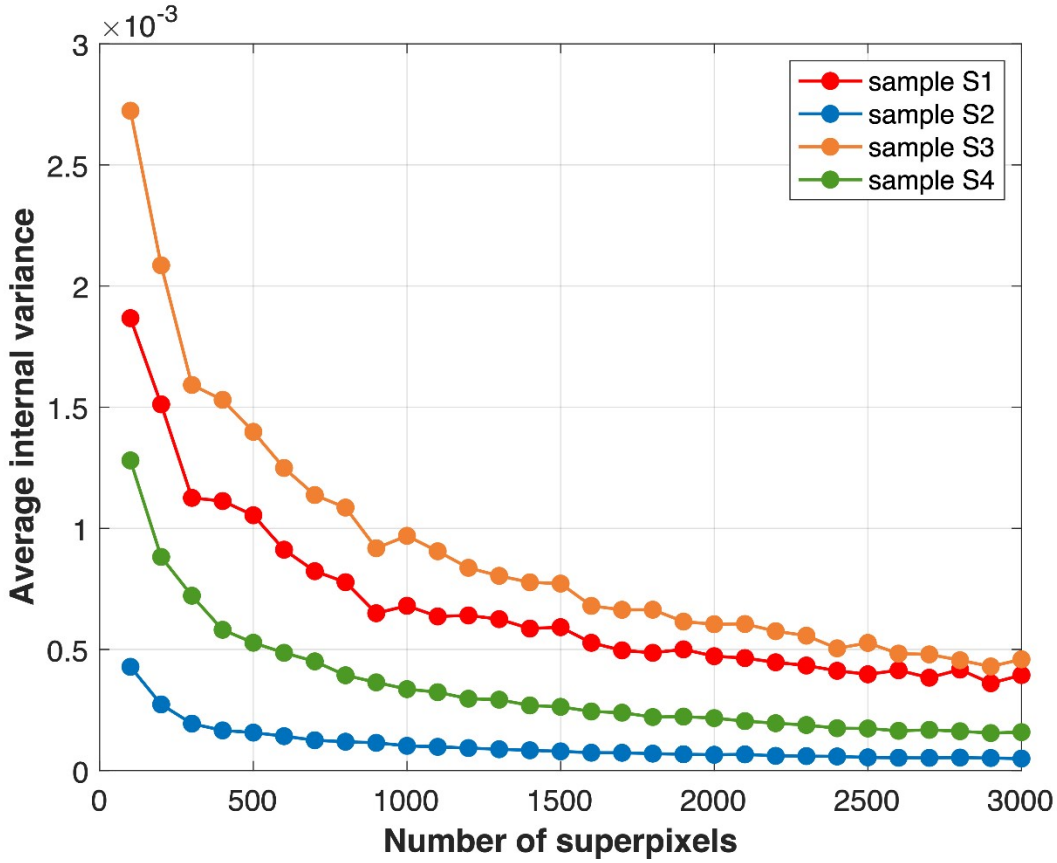

Fig.SM2 Graphical representation of the average internal variance of the superpixels calculated for each frame as a function of the number of superpixels. The curve stabilizes at around 2000 superpixels, indicating that the internal variance has become nearly homogeneous.

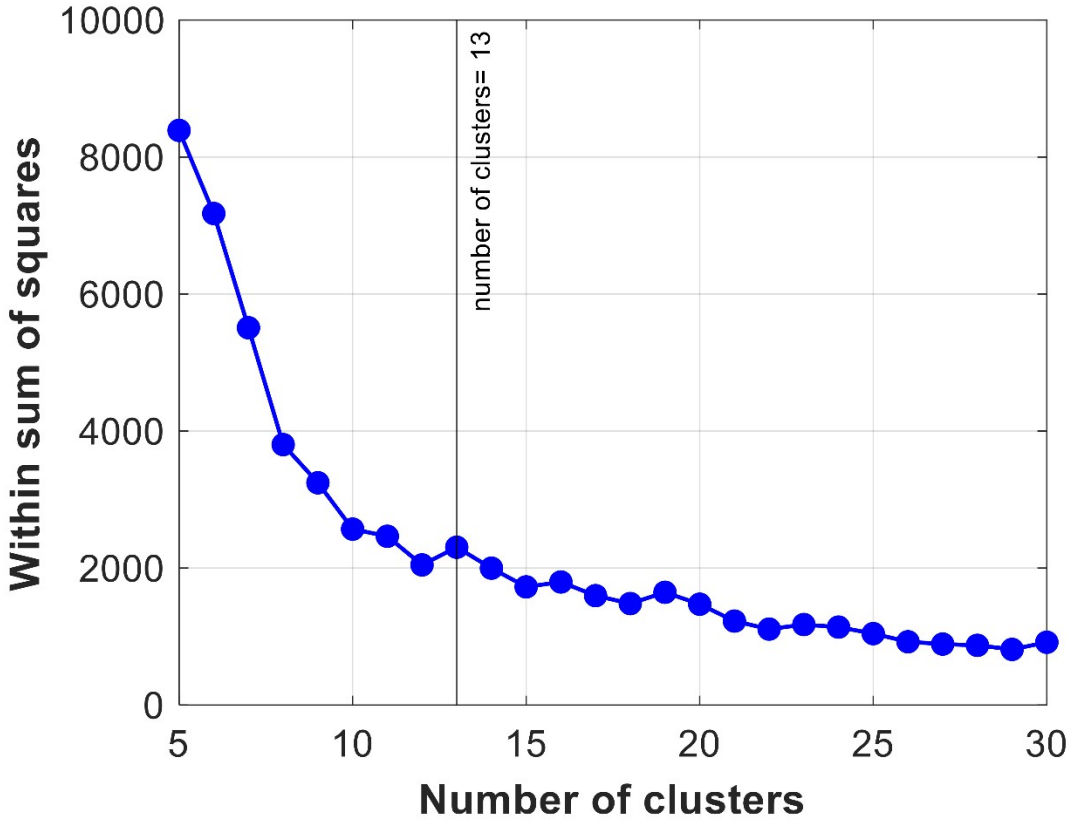

Fig.SM3 the Within-Cluster Sum of Square (WCSS) values plotted against the different number of clusters computed using the GMM. A flattening in the WCSS curve is observed between 10 and 13 clusters. 13 clusters were selected to provide a more faithful representation of the color variations.

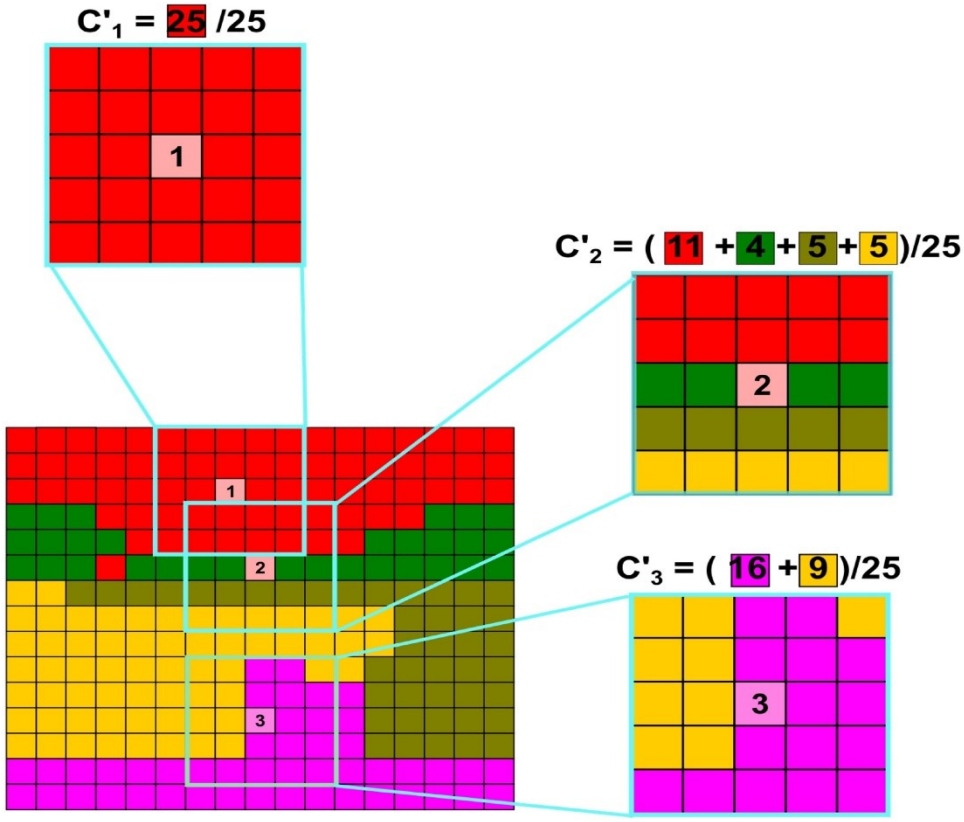

Fig.SM4 Graphical representation for the calculation of the weights, used to remove the border effects of the clusters. For each pixel, the weights are computed counting the number of pixels belonging to each cluster in a selected neighborhood. This approach ensures a smooth color transition across cluster boundaries.

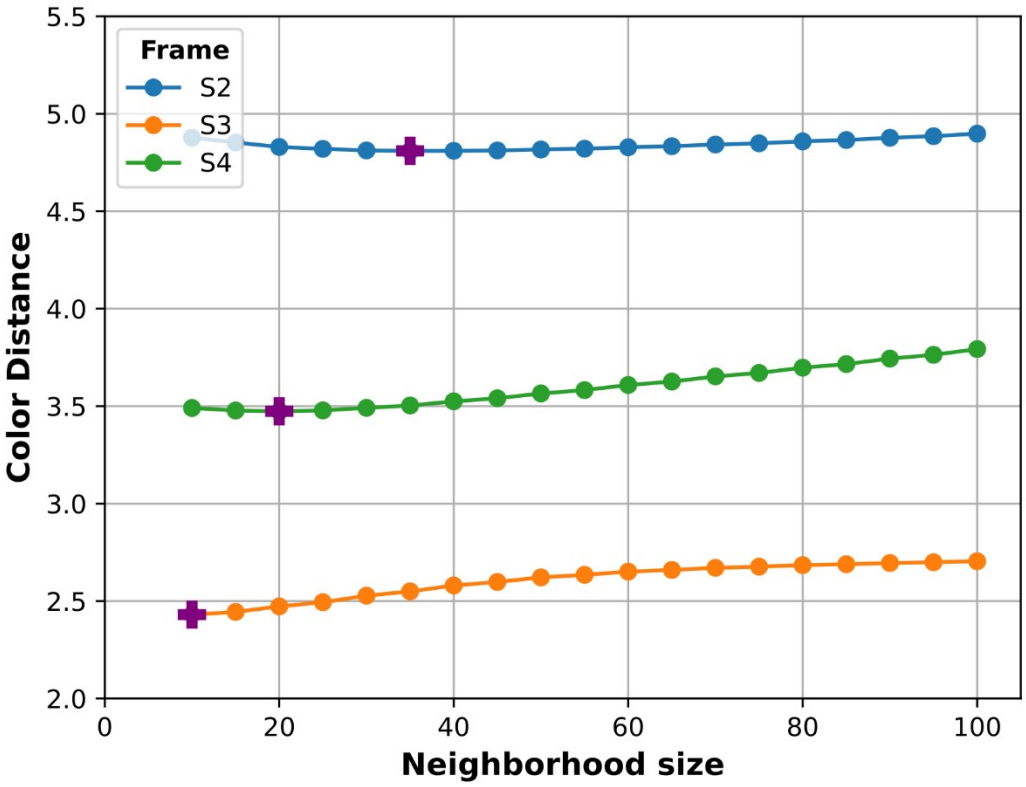

Fig. SM5 Plot showing the color distance between the corrected frames and the reference one at different neighborhood size values. The purple cross represents the minimum distance between the images, consequently the optimal number of neighborhood pixels.