

Correlation-based Alignment of Raw Endoscopic Sequence Data with Physician Selected Movies

Workshop Germany Brazil 2016

Understanding the aggressiveness of cancer cells through novel imaging techniques

Marc Aubreville¹, Christian Jaremenko¹, Christian Knipfer^{2,3}, Nicolai Oetter^{2,4},
Helmut Neumann^{2,5}, Florian Stelzle^{2,3} and Andreas Maier^{1,2}

¹ Pattern Recognition Lab, Computer Science 5, Friedrich-Alexander University Erlangen-Nürnberg (FAU)

² Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander University Erlangen-Nürnberg (FAU)

³ Department of Oral and Maxillofacial Surgery, University Medical Center Hamburg-Eppendorf

⁴ Department of Oral and Maxillofacial Surgery, Friedrich-Alexander University Erlangen-Nürnberg (FAU), University Hospital Erlangen

⁵ First Department of Internal Medicine, University Medical Center, Johannes Gutenberg University Mainz

marc.aubreville@fau.de

We present a simple yet robust algorithm to register low bit-resolution medical video data to high resolution raw image sequences. The approach was tested using confocal laser microendoscopy images and yielded better results than a cross correlation approach on a single image.

Keywords: *high-definition range image alignment, confocal laser microendoscopy*

I. BACKGROUND

Medical imaging data is often recorded in high bit resolutions, as little changes in contrast may be essential for diagnosis of pathologies. This is especially true for contrast-agent based confocal laser microendoscopy images. For displaying purposes, a dynamic mapping from those high-definition range images to an 8-bit grayscale is performed. After the examination session, characteristic sequences are exported in a low-definition video format, by applying the same (frame dependent) mapping. For automatic detection of cell borders [1,2] or classification of malignant tissue [3], it is however important to have unmodified data, to be independent from the gray value scaling applied.

II. AIMS AND METHODS

The objective of the present work is thus to find for given 8-bit movie images L_n the corresponding indices in a raw image sequence H_m for later automatic classification of those images using machine learning techniques.

We used raw data in Cellvizio MKT format (Cellvizio, Mauna Kea Technologies, Paris, France) that was imported using a custom framework for image data [4] to retrieve images of 16-bit grayscale data. Since exported video in our case is annotated with a logo and a dimension legend (see figure 1), we apply a circular mask as a preprocessing step. In order to find the correct indices of the N frames selected in

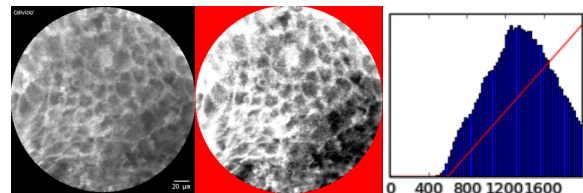


Figure 1: Original CLE video (left), scaled raw video (middle) with circular mask (red) and gray level histogram (right) with scaling (red line)

the low bit-resolution video file, Pearson's correlation was computed between the first video image, the last video image and each of the M raw images:

$$\begin{bmatrix} r_{1,1} \cdot r_{1,M} \\ r_{2,1} \cdot r_{2,M} \end{bmatrix} = \begin{bmatrix} \frac{COV(L_0, H)^T}{\sigma_L \sigma_H}, \frac{COV(L_N, H)^T}{\sigma_N \sigma_H} \end{bmatrix}^T$$

This results in two vectors of correlation values. For many images and sequences, it is sufficient to evaluate the maxima in the first vector in order to find the respective offset of the sequence. For other sequences, however, this approach fails (see figure 2, left). In order to make the algorithm more robust, we incorporate the known video length N :

$$i_{start} = \operatorname{argmax}_{i \in [0, M-N]} (r_{i,1} \cdot r_{i+N-1,2})$$

The criterion effectively multiplies the vectors with the video length N as an offset, yielding maximum values if high values occur in both vectors N samples apart. The resulting i_{start} is the frame index of the

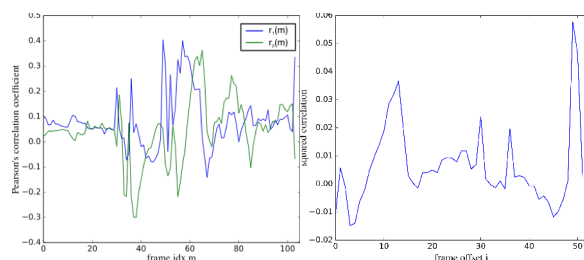


Figure 2: Correlation coefficient for pictures (left) and proposed metric (product of shifted correlations) (right).

first frame of the low-definition movie within the high-definition movie.

Considering the example of figure 2, the proposed metric yields a clear maximum (right side of figure), corresponding to the correct offset.

III. RESULTS AND CONCLUSION

Visual evaluation has shown that the method has proven to provide perfect matching on a confocal laser microscopy image set [5], consisting of 96 sequences, where a simple correlation of the first frame was erroneous in some (n=8) cases.

IV. REFERENCES

- [1] Mualla, F., Schöll, S., Bohr, C., Neumann, H., & Maier, A. (2014). Epithelial Cell Detection in Endomicroscopy Images of the Vocal Folds. In *International Multidisciplinary Microscopy Congress (Vol. 154, pp. 201–205)*. Cham: Springer International Publishing.
- [2] Bier, B., Mualla, F., Steidl, S., Bohr, C., Neumann, H., Maier, A., & Hornegger, J. (2015). Band-Pass Filter Design by Segmentation in Frequency Domain for Detection of Epithelial Cells in Endomicroscope Images. In *Bildverarbeitung für die Medizin 2015 (pp. 413–418)*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- [3] Jaremenko, C., Maier, A., Steidl, S., Hornegger, J., Oetter, N., Knipfer, C., et al. (2015). Classification of Confocal Laser Endomicroscopic Images of the Oral Cavity to Distinguish Pathological from Healthy Tissue. In H. Handels, T. M. Deserno, H.-P. Meinzer, & T. Tolxdorff (Eds.), *Bildverarbeitung für die Medizin 2015 (pp. 479–485)*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- [4] Framework for CLE sequence files, published on github: <https://github.com/maubreville/CLEtools>
- [5] Oetter, N., Knipfer, C., Rohde, M., Wilmowsky, C., Maier, A., Brunner, K., et al. (2016). Development and validation of a classification and scoring system for the diagnosis of oral squamous cell carcinomas through confocal laser endomicroscopy. *Journal of Translational Medicine*, 1–11.