Super-Resolved Retinal Image Mosaicing

2016 IEEE International Symposium on Biomedical Imaging (ISBI'16)

Thomas Köhler, Axel Heinrich, Andreas Maier, Joachim Hornegger, Ralf Tornow
15.04.2016
Pattern Recognition Lab, Friedrich-Alexander-Universität Erlangen-Nürnberg
Erlangen Graduate School in Advanced Optical Technologies (SAOT)
Departm. of Ophthalmology, Friedrich-Alexander-Universität Erlangen-Nürnberg

SAOT erlangen graduate school in advanced optical technologies

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Introduction

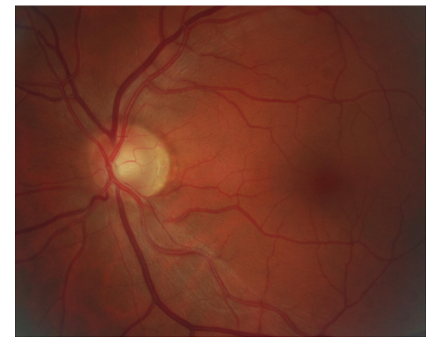
erlangen
graduate school in
advanced optical technologies

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Background

Structural imaging technologies in ophthalmology

- Optical coherence tomography (3-D)
- Slit lamp examination (2-D)
- Scanning laser ophthalmoscopy (2-D)
- **Fundus photography / video imaging (2-D / 2-D + t)**



Non-invasive examination of the eye background

- For diagnosis or screening
- For interventional applications

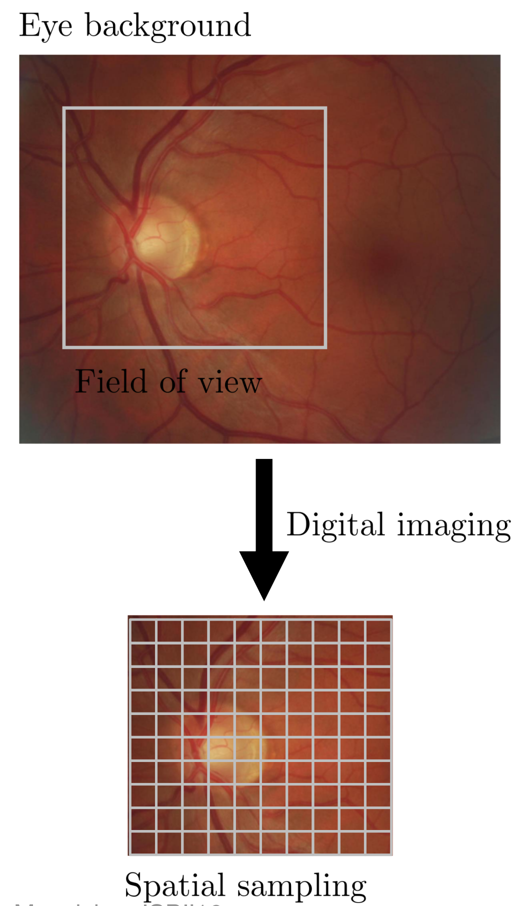
Background

Characteristics of today's imaging systems

- Narrow field of view (20 – 50°)
  - Limited by the optics and the anatomy
  - Pupil dilation for large field of view
- Limited spatial resolution
  - Limited by the sensor array and the optics
  - High resolution desirable to examine small anatomical structures

Solutions to overcome the limitations

- Hardware-based ⇒ expensive and/or not mobile (limitations for screening applications)
- **Software-based ⇒ low-cost solution**



Related Work on Software-Based Methods

- Mosaicing methods based on image registration
  - Feature-based registration[1]
  - Intensity-based registration[2]

  ⇒ Enhance the field of view at the same spatial resolution from multiple images acquired longitudinally

- Multi-frame super-resolution[3]

  ⇒ Enhance the spatial resolution at the same field of view from multiple images acquired during one examination

Mosaicing and super-resolution are complementary techniques

[1] A. Can et al., *A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina*, IEEE PAMI, vol. 24, no. 3, pp. 347–364, 2002.
[2] K. M. Adal et al., *A Hierarchical Coarse-to-Fine Approach for Fundus Image Registration*, Proc. WBIR 2014, 2014, pp. 93–102
[3] D. Thapa et al., *Comparison of superresolution algorithms applied to retinal images*, Journal of Biomedical Optics, vol. 19, no. 5, pp. 056002, 2014.

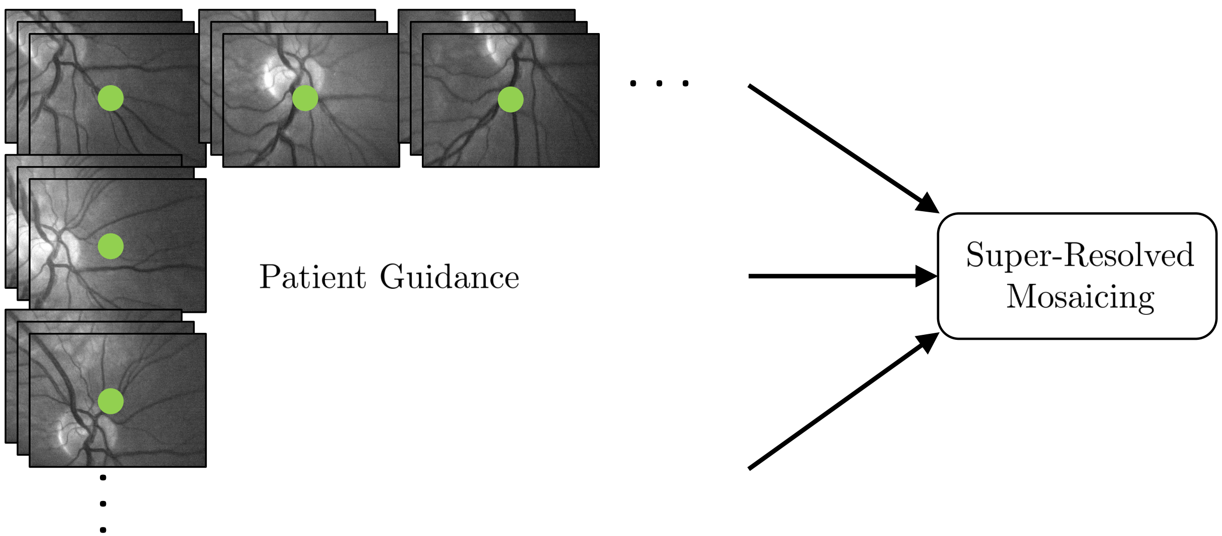
Super-Resolved Mosaicing Framework

SAOT
erlangen
graduate school in
advanced optical technologies

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Key Idea and Acquisition Protocol

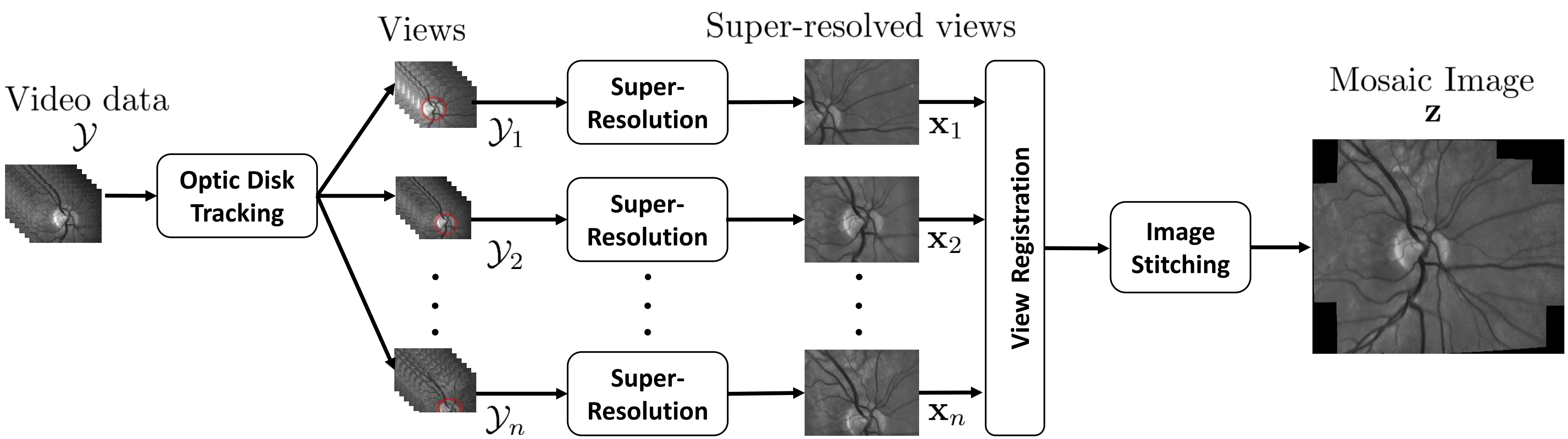
- Patient guidance to control viewing direction (e. g. fixation target)
  ⇒ Scan different regions on the retina
- Acquisition of video sequence over the entire examination



Reconstruct image with enhanced field of view and spatial resolution from low-resolution video

Framework Overview

Given: K frames $\mathcal{Y} = \left\{\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(K)}\right\}$ that show different regions on the retina

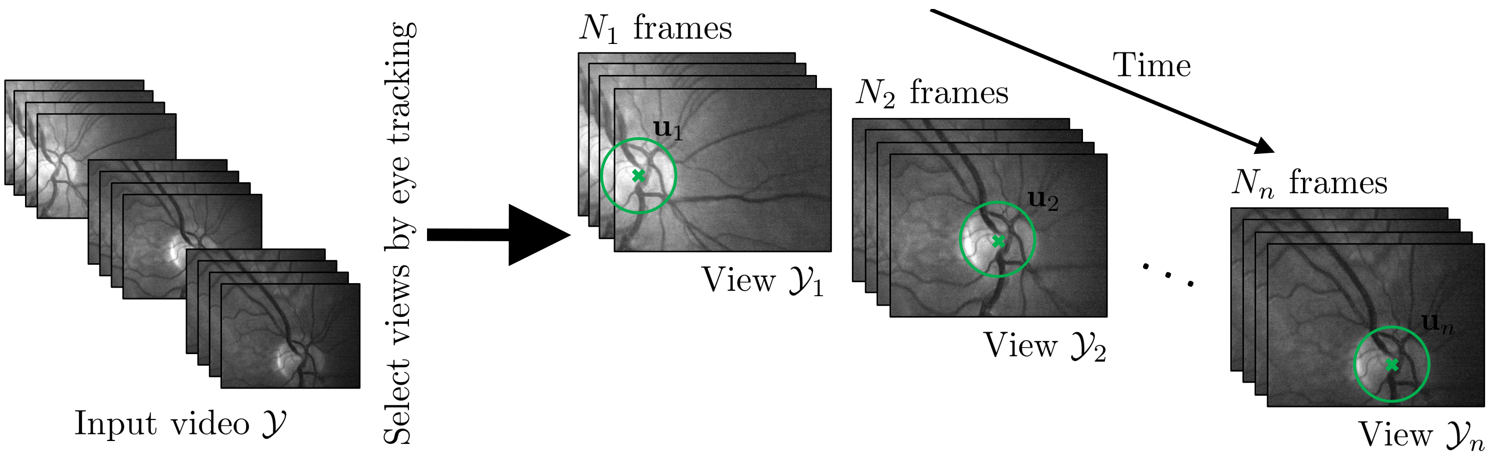


Three-stage approach to super-resolved mosaicing:

1. Select n views $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$ from $\mathcal{Y}$ by means of eye tracking
2. Combine low-resolution frames in each view $\mathcal{Y}_i$ to a super-resolved image $\boldsymbol{x}_i$
3. Combine super-resolved images $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ to the common mosaic $\boldsymbol{z}$

Automatic View Selection

- Track the eye position over time using the optic disk center as a feature
- Fully automatic in real-time using tracking-by-detection scheme[4]
- Select frames for each view $\mathcal{Y}_i$ according to:
  - Large eye motion **due to patient guidance** between successive views ($d(\boldsymbol{u}_{i-1}, \boldsymbol{u}_i) \geq d_{\min}$)
  - Small eye motion **due to saccades** between successive frames within one view



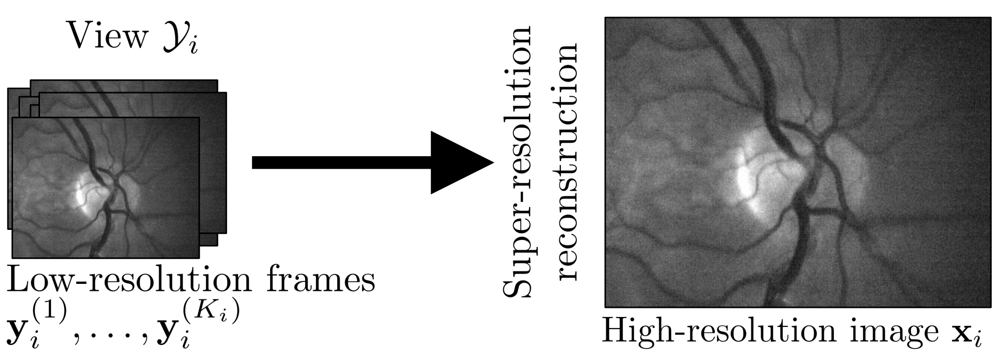
[4] A. Kürten et al., *Geometry-based optic disk tracking in retinal fundus videos*, Proc. BVM 2014, pp. 120–125. 2014.

Super-Resolution View Reconstruction

- Given K_i low-resolution frames $\boldsymbol{y}_i^{(1)}, \ldots, \boldsymbol{y}_i^{(K_i)}$ for view $\mathcal{Y}_i$
- Reconstruct high-res. image $\boldsymbol{x}_i$ via maximum a-posteriori estimation:

$$\boldsymbol{x}_i = \arg\min_{\boldsymbol{x}} \sum_{k=1}^{K_i} \underbrace{\left|\left| \boldsymbol{y}_i^{(k)} - \gamma_{m,i}^{(k)} \odot \boldsymbol{W}_i^{(k)}\left(\boldsymbol{\theta}_i^{(k)}\right)\boldsymbol{x} - \gamma_{a,i}^{(k)}\mathbf{1} \right|\right|_1}_{\text{geometric model } (\boldsymbol{\theta}_i^{(k)}) \text{ + photometric model } (\gamma_{m,i}^{(k)} \text{ and } \gamma_{a,i}^{(k)})} + \underbrace{\lambda(\boldsymbol{x}) \cdot R(\boldsymbol{x})}_{\text{bilateral total variation}}$$

  - Determine geometric and photometric parameters by pair-wise registration
  - Determine regularization weight $\lambda(\boldsymbol{x})$ by image quality self-assessment[5]



[5] T. Köhler et al., *Multi-frame Super-resolution with Quality Self-assessment for Retinal Fundus Videos*, MICCAI 2014, 2014, pp. 650–657

Intensity-Based View Registration

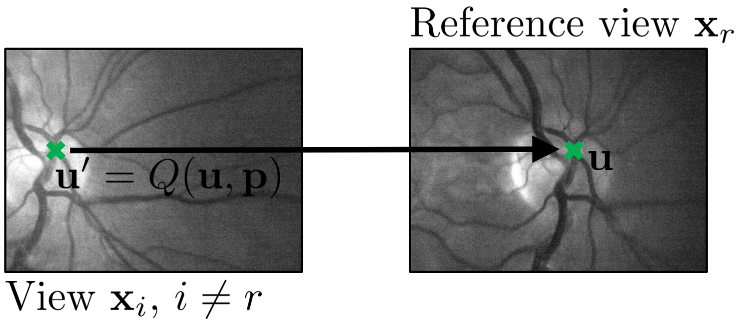
- Quadratic transformation $\boldsymbol{u}' = Q(\boldsymbol{u}, \mathbf{p})$ for point $\boldsymbol{u}$ in a reference view to point $\boldsymbol{u}'$:

$$\boldsymbol{u}' = \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 \\ p_7 & p_8 & p_9 & p_{10} & p_{11} & p_{12} \end{pmatrix} \begin{pmatrix} u_1^2 & u_2^2 & u_1 u_2 & u_1 & u_2 & 1 \end{pmatrix}^\top$$

- Correlation based similarity measure $\rho : \mathbb{R}^N \times \mathbb{R}^N \rightarrow [-1; 1]$:

$$\hat{\boldsymbol{p}} = \arg\max_{\boldsymbol{p}} \rho \left\{ \boldsymbol{x}_r(\boldsymbol{u}), \boldsymbol{x}_i(Q(\boldsymbol{u}, \mathbf{p})) \right\}$$

- Enhanced correlation coefficient (ECC) maximization[6] with quadratic transformation model $\Rightarrow$ iterative optimization of ρ in coarse-to-fine scheme



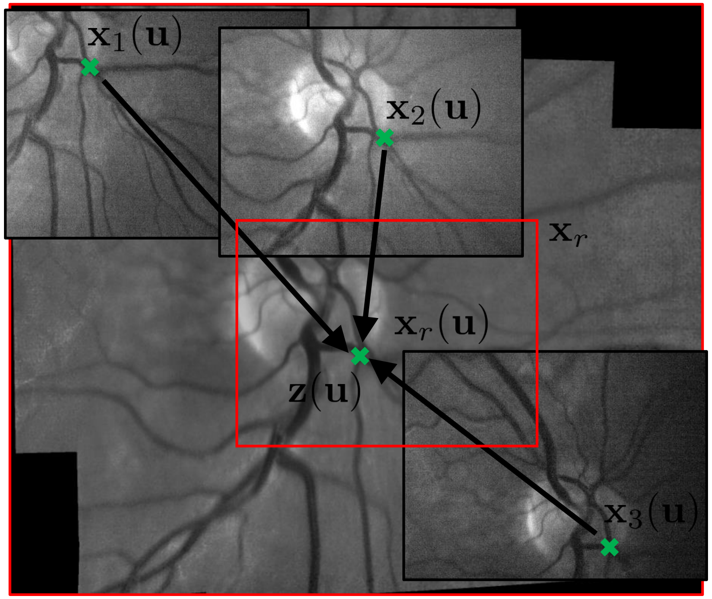
[6] G. D. Evangelidis and E. Z. Psarakis, *Parametric image alignment using enhanced correlation coefficient maximization.*, IEEE PAMI, vol. 30, no. 10, pp. 1858–65, 2008.

Mosaic Image Reconstruction

- Warp super-resolved views $\boldsymbol{x}_i$, $i \neq r$ towards the reference view $\boldsymbol{x}_r$
- Histogram matching for illumination correction $\Rightarrow$ registered views $\tilde{\boldsymbol{x}}_1, \ldots, \tilde{\boldsymbol{x}}_n$
- Stitching by pixel-wise adaptive averaging:

$$\boldsymbol{z}(\boldsymbol{u}) = \frac{1}{\sum_{i=1}^{n(\boldsymbol{u})} \boldsymbol{w}_i(\boldsymbol{u})} \sum_{i=1}^{n} \boldsymbol{w}_i(\boldsymbol{u}) \tilde{\boldsymbol{x}}_i(\boldsymbol{u})$$

$\boldsymbol{w}_i(\boldsymbol{u})$: adaptive weight at pixel position $\boldsymbol{u}$ associated with the i-th view



Experiments and Results

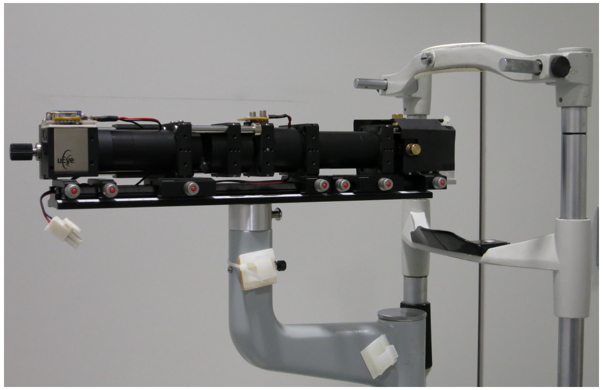
SAOT
erlangen
graduate school in
advanced optical technologies

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Experimental Setup

Video acquisition with low-cost fundus camera[7]

- Monochromatic video (25 Hz)
- VGA resolution (640 × 480 px) with FOV of 15° in vertical and 20° in horizontal direction



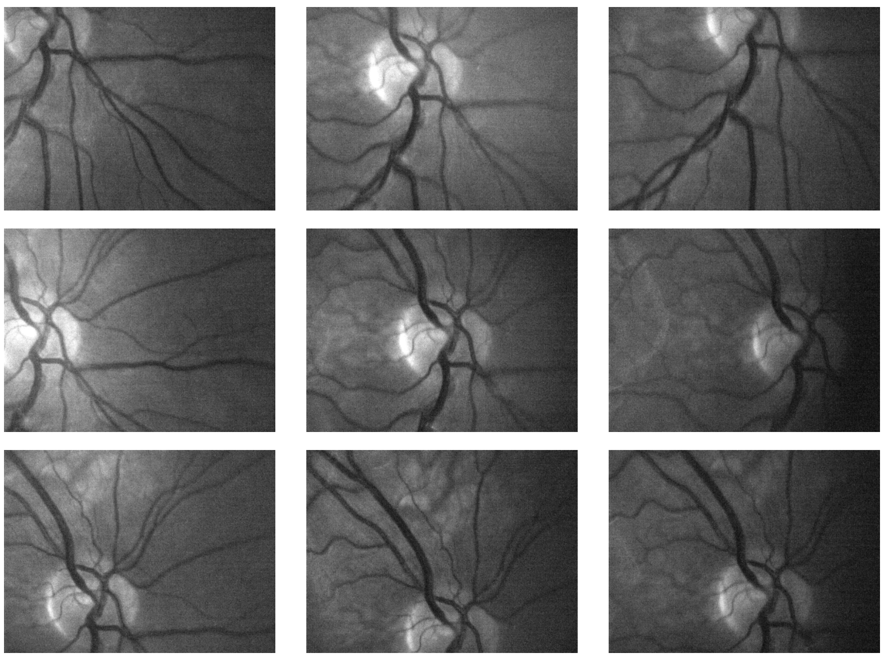
Evaluation data

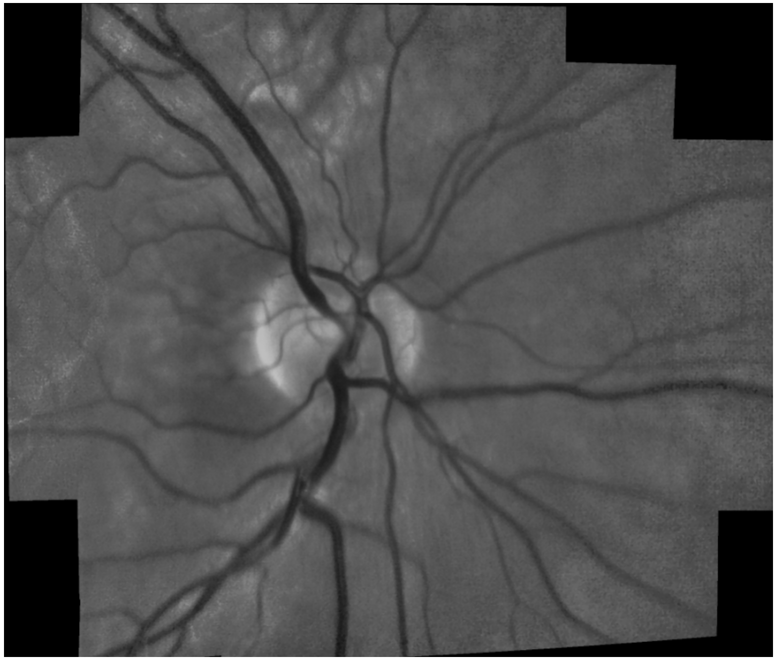
- Examination of seven human subjects (left eye) without pupil dilation ⇒ 24 datasets
- Horizontal/vertical eye movements during examination ⇒ 3 to 9 views for each subject
- ≈ 15 s video per examination

[7] R. P. Tornow et al., *Non-mydriatic video ophthalmoscope to measure fast temporal changes of the human retina*, Proc. SPIE Novel Biophotonics Techniques and Applications, 2015, vol. 9540, pp. 954006–954006–6

Qualitative Examples

Input frames from $n = 9$ views and super-resolved mosaic image



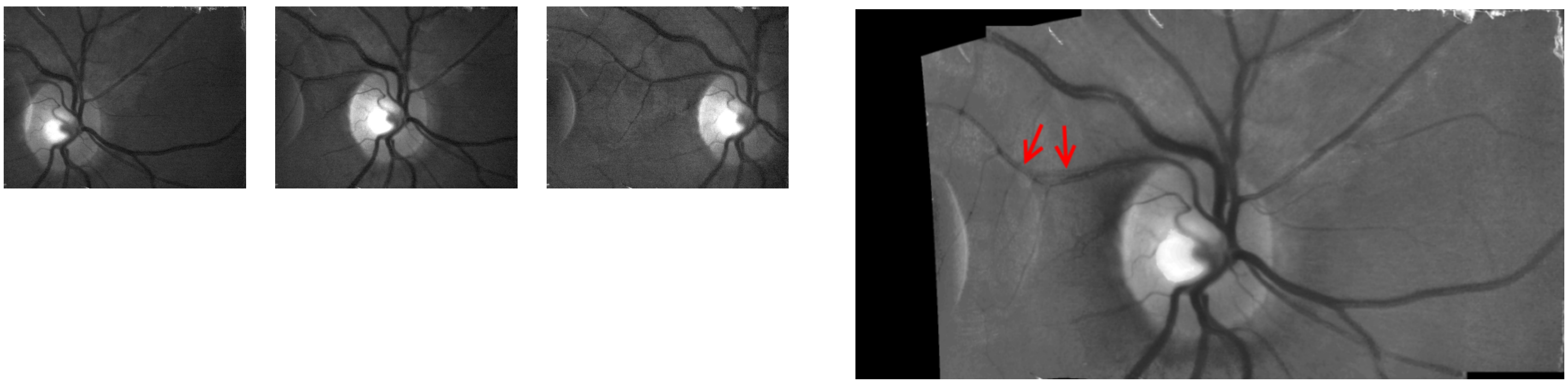


$\Rightarrow$ Increased spatial sampling by a factor of two

$\Rightarrow$ Increased field of view from 15° to 25°

Qualitative Examples

Input frames from $n = 3$ views and super-resolved mosaic image



⇒ Local misregistration between left and central view in the final mosaic
⇒ But no error accumulation due to registration with fixed reference

Objective Image Noise / Sharpness Evaluation

- Blind signal-to-noise ratio (SNR) estimation in homogeneous regions:

$$Q_{\text{snr}} = 10 \log_{10} \left(\frac{\mu_{\text{flat}}}{\sigma_{\text{flat}}} \right)$$

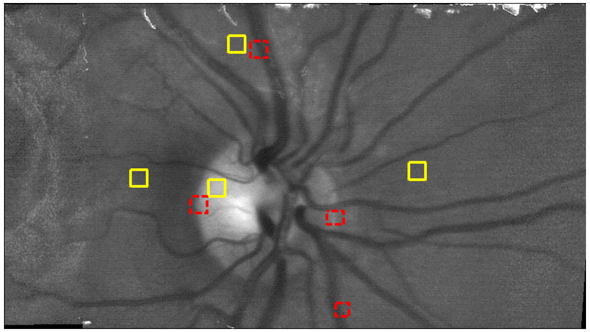
  μ_{flat} and σ_{flat} denote mean / standard deviation in a homogenous region
- Blind edge preservation measurement in regions containing an object boundary:

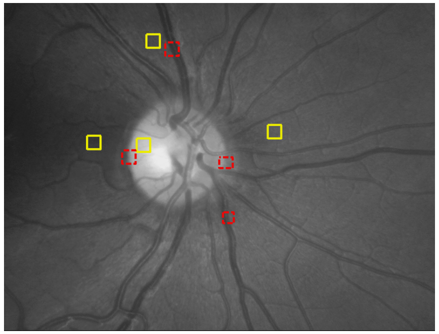
$$Q_{\text{edge}} = \frac{w_b(\mu_b - \mu)^2 + w_f(\mu_f - \mu)^2}{w_b\sigma_b^2 - w_f\sigma_f^2}$$

  w_i, μ_i and σ_i with $i \in \{b, f\}$ denote the weight, the mean and the standard deviation of a Gaussian mixture model fitted for background (b) and foreground (f)

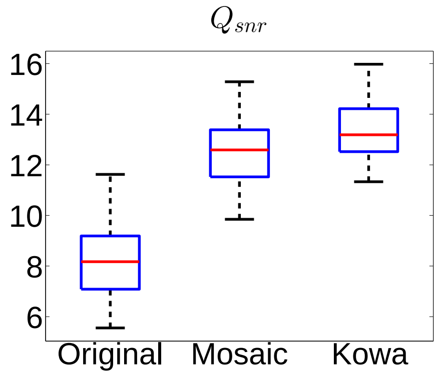
Objective Image Noise / Sharpness Evaluation

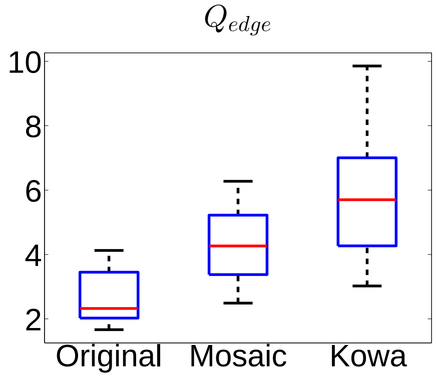
- Super-resolved mosaic and photograph of Kowa nonmyd camera (25° FOV, 1600×1216 px, grayscale-converted) with regions of interests:





- Statistics of Q_{snr} and Q_{edge} for five human subjects:





Summary and Conclusion

erlangen graduate school in advanced optical technologies

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

TECHNISCHE FAKULTÄT

Summary and Conclusion

Summary

- Fully automatic approach to super-resolved mosaicing
- Reconstruction of high-resolution retinal images of enlarged field of view from low-resolution video
- Applicable to low-cost retinal fundus imaging

Future work

- **Algorithm:** Super-resolution and mosaicing as joint optimization problem
  $\Rightarrow$ Further improve robustness of mosaicing
- **Applications:** Investigations of clinical applications
  $\Rightarrow$ e. g. screening for eye diseases

Thank you very much for your attention!