MICCAI 2014 BOSTON

Multi-Frame Super-Resolution with Quality Self-Assessment for Retinal Fundus Videos

T. Köhler[1,2], A. Brost[1], K. Mogalle[1], Q. Zhang[1], C. Köhler[3], G. Michelson[2,3], J. Hornegger[1,2], R. P. Tornow[3]

[1] Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg
[2] Erlangen Graduate School in Advanced Optical Technologies (SAOT)
[3] Department of Ophthalmology, Friedrich-Alexander-Universität Erlangen-Nürnberg
thomas.koehler@fau.de

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Background

Fundus video cameras enable the **non-invasive acquisition of fast temporal changes** on the human retina. The low spatial resolution limits their diagnostic applicability.

Contribution: Reconstruction of high-resolution retinal fundus images from low-resolution, low-cost video data by means of **multi-frame super-resolution** to enhance diagnostic usability.

Problem Statement

Generative image model: Each frame $\boldsymbol{y}^{(k)}$ for $k = 1, \ldots, K$ is a low-resolution observation of the high-resolution image $\boldsymbol{x}$:

$$\begin{aligned}\boldsymbol{y}^{(k)} &= \boldsymbol{\gamma}_m^{(k)} \odot \boldsymbol{D}\boldsymbol{B}^{(k)}\boldsymbol{M}^{(k)}\boldsymbol{x} + \gamma_a^{(k)}\mathbf{1} + \boldsymbol{\epsilon}^{(k)} \\ &= \boldsymbol{\gamma}_m^{(k)} \odot \boldsymbol{W}^{(k)}\boldsymbol{x} + \gamma_a^{(k)}\mathbf{1} + \boldsymbol{\epsilon}^{(k)}\end{aligned}$$

- Eye motion modeled by geometric transformation $\boldsymbol{M}^{(k)}$
- Point spread function (PSF) $\boldsymbol{B}^{(k)}$, subsampling $\boldsymbol{D}$, noise $\boldsymbol{\epsilon}^{(k)}$
- Inhomogeneous illumination modeled by $\boldsymbol{\gamma}_m^{(k)}$ and $\gamma_a^{(k)}$

Objective: Given a sequence of K low-resolution frames $\boldsymbol{y}^{(1)}, \ldots, \boldsymbol{y}^{(K)}$, reconstruct unknown high-resolution image $\boldsymbol{x}$

Proposed Super-Resolution Framework

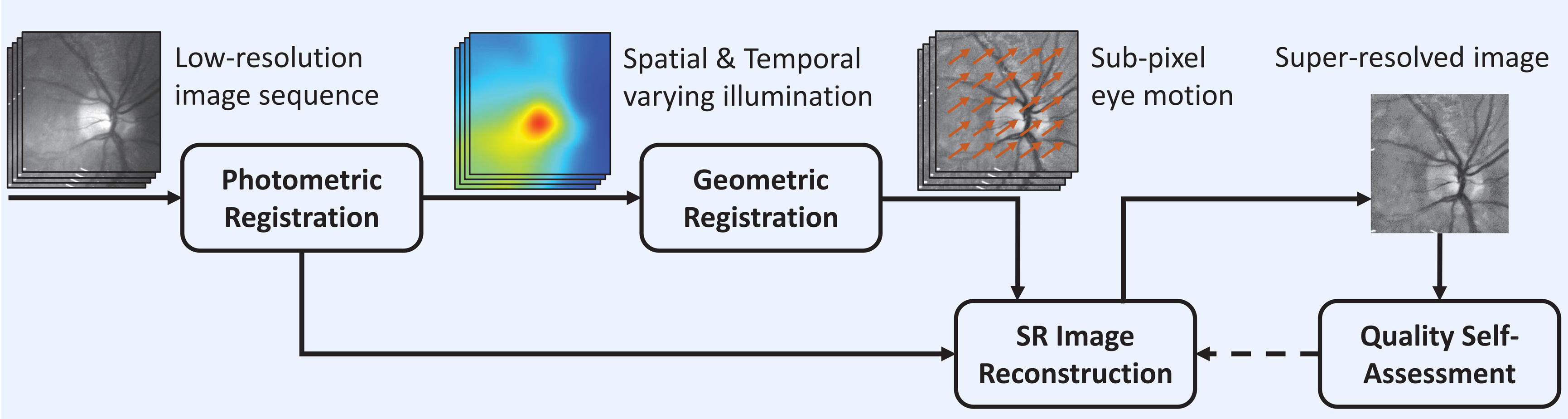


Figure 1: Flowchart of the proposed multi-frame super-resolution framework.

1. **Photometric and geometric registration** for each low-resolution frame $k = 2, \ldots, K$ using $k = 1$ as reference:
   - Estimate photometric parameter $\boldsymbol{\gamma}_m^{(k)}$ using B-spline approximation [1] of the bias field and $\gamma_a^{(k)}$ by the median temporal brightness difference
   - Small, natural eye motion is modeled by $\boldsymbol{M}^{(k)}$ and subpixel motion is estimated by affine registration [2]
2. **Super-resolution image reconstruction:**

   $$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \left\{ \sum_{k=1}^{K} \left\| \boldsymbol{y}^{(k)} - \boldsymbol{\gamma}_m^{(k)} \odot \boldsymbol{W}^{(k)}\boldsymbol{x} - \gamma_a^{(k)}\mathbf{1} \right\|_1 + \lambda R(\boldsymbol{x}) \right\}$$

   $R(\boldsymbol{x})$ is an edge preserving regularizer weighted by λ
   $\rightarrow$ Iterative scaled conjugate gradient (SCG) optimization
3. **Image quality self-assessment:**
   - Quality metric $Q_v(\boldsymbol{x})$ for anisotropic patches $\boldsymbol{p}_i \in \mathcal{P}(\boldsymbol{x})$:

     $$Q_v(\boldsymbol{x}) = \sum_{\boldsymbol{p}_i \in \mathcal{P}(\boldsymbol{x})} \sigma_i \cdot q(\boldsymbol{p}_i)$$

     Local quality score $q(\boldsymbol{p}_i)$ adaptively weighted by local variance σ_i estimated from of the vesselness in $\boldsymbol{p}_i$ [3]
   - Automatic selection of regularization weight λ:

     $$\hat{\lambda} = \arg\max_{\lambda} Q_v(\boldsymbol{x}_\lambda)$$

     $\boldsymbol{x}_\lambda$: super-resolved image reconstructed with weight λ
     $\rightarrow$ Joint super-resolution and parameter selection

Experiments and Results

1. **Experiments for synthetic fundus images:**
   - Ground truth images taken from the DRIVE database
   - Simulation: random translations for eye motion, blurring by a Gaussian PSF, subsampling and Gaussian noise
   - Super-resolution: $K = 16$ frames for $2\times$ magnification

   Evaluation of the peak-signal-to-noise ratio (PSNR) and influence of super-resolution to blood vessel segmentation:

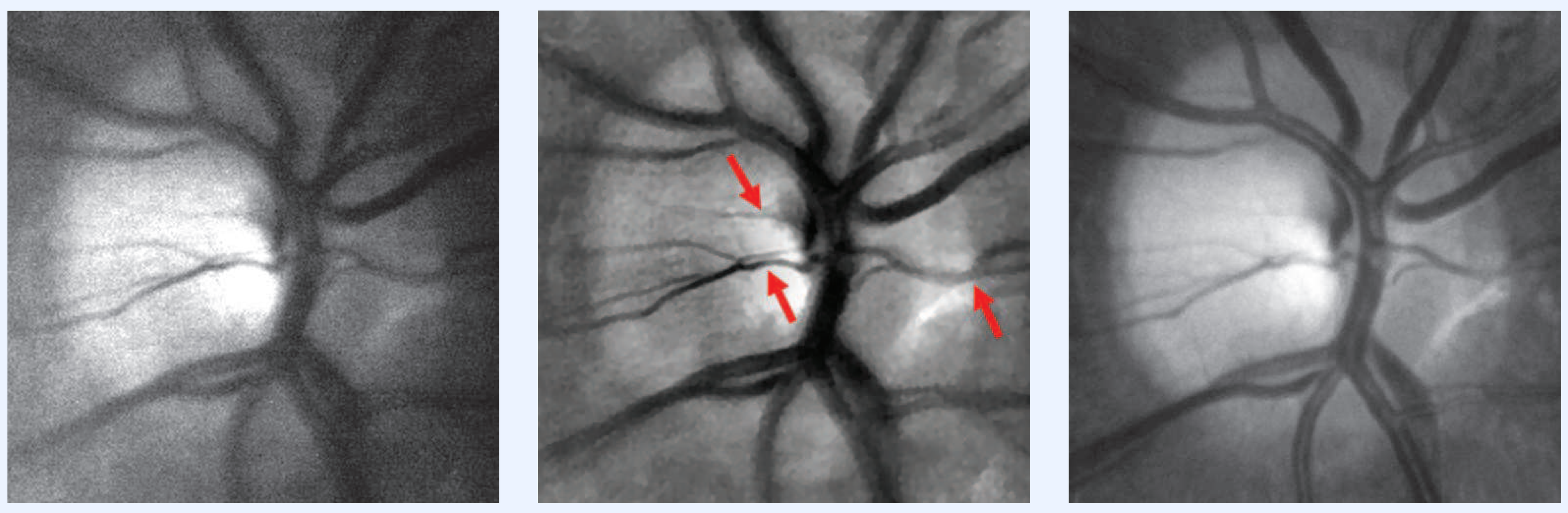
	LR frame	SR image	Ground truth
PSNR (in dB)	31.09 $\pm$ 3.10	**31.92 $\pm$ 3.39**	-
Sensitivity (%)	57.59 $\pm$ 6.01	**70.37 $\pm$ 5.00**	72.85 $\pm$ 6.70
Specificity (%)	94.31 $\pm$ 1.40	**93.99 $\pm$ 1.26**	94.57 $\pm$ 1.34

Figure 2: Peak-signal-to-noise ratio (PSNR) for synthetic images generated from the DRIVE database as well as sensitivity and specificity for automatic vessel segmentation.

   Super-resolution enhanced mean PSNR by 0.8 dB and sensitivity for automatic vessel segmentation by 13 %

2. **Experiments for low-cost fundus video camera:**
   - Videos of 6 healthy subjects ($\approx$ 15 s video per subject)
   - System parameters: CCD camera (640$\times$480 px), 20° field of view (FOV), 12.5 Hz frame rate
   - Super-resolution: $K = 8$ frames for $2\times$ magnification

   Comparison of super-resolved video data to high-resolution Kowa NonMyd images (1600$\times$1216 px, 25° FOV):



(a) Original (low-resolution) frame (b) Super-resolved image (c) Kowa NonMyd image

Figure 3: Low-resolution and super-resolved image (magnification factor: 2) of the optic nerve head in comparison to a high-resolution image acquired with a Kowa NonMyd camera (images contrast enhanced for visual comparison).

   Q_v measure improved by 160 % w.r.t. to low-resolution data (compared to 86 % for temporal median filtering)

Conclusion

Novel super-resolution framework for fundus video imaging:

- **Geometric and photometric registration** to model eye motion and inhomogeneous illumination
- **Image quality self-assessment** for automatic selection of regularization parameter
- Quality of **super-resolved low-cost images** comparable to a high-resolution Kowa NonMyd camera

Acknowledgements

The authors gratefully acknowledge funding of the Erlangen Graduate School in Advanced Optical Technologies (SAOT) by the German National Science Foundation (DFG) in the framework of the excellence initiative.

References

[1] Radim Kolar, Jan Odstrcilik, Jiri Jan, and Vratislav Harabis. Illumination Correction and Contrast Equalization in Colour Fundus Images. In *Proceedings EUSIPCO 2011*, pages 298–302, 2011.

[2] Georgios D. Evangelidis and Emmanouil Z. Psarakis. Parametric image alignment using enhanced correlation coefficient maximization. *IEEE Transactions on Pattern Analysis Machine Intelligence*, 30(10):1858–65, 2008.

[3] Thomas Köhler, Attila Budai, Martin F. Kraus, Jan Odstrcilik, Georg Michelson, and Joachim Hornegger. Automatic no-reference quality assessment for retinal fundus images using vessel segmentation. In *Proceedings CBMS 2013*, pages 95–100, 2013.