Image Fusion Through Linear Embeddings

Oguzhan Ulucan, Diclehan Karakaya, Mehmet Turkan

Department of Electrical and Electronics Engineering, Izmir University of Economics, Izmir, Turkey

September 2021

Presenter: Oguzhan Ulucan Email: oguzhanulucan.de@gmail.com Check out our paper @1085

Content

- Background
- Proposed Method
  - Determination of main exposures
  - Weight maps via linear embeddings
  - Adaptive morphological masking
  - Exposure fusion and post-processing
- Experimental Results
  - Application to visible and infrared image fusion
- Conclusion

Background

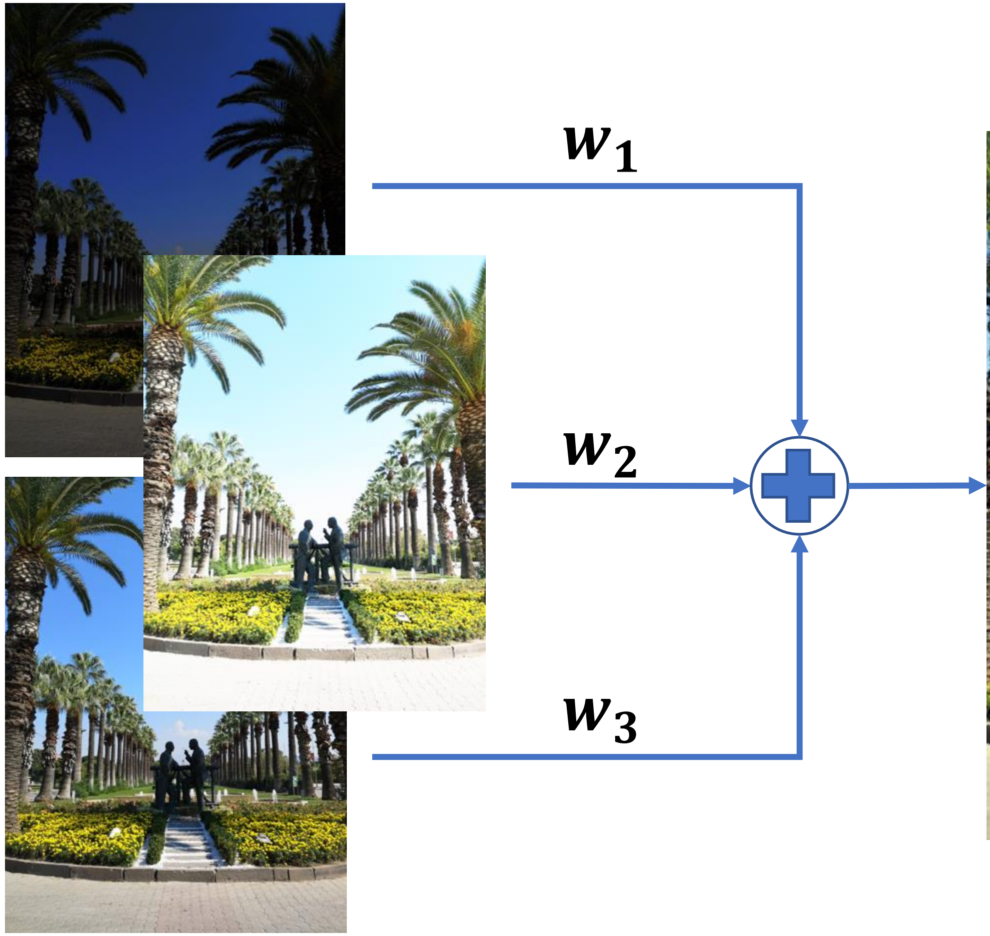


Figure. LDR** Input Stack, *Izmir Fair*.

Figure. HDR-like Image.

Aim

- Create HDR*-like content
  - Fine details
  - Vivid colors
- By taking advantage of
  - Image Fusion
  - Linear Embeddings
  - Morphological Masking

* High dynamic range
** Low dynamic range

LDR Input stack, *Izmir Fair*, courtesy of Erdem Okur.

Background

- High contrast scenery

  Brightest pixel intensity – Darkest pixel intensity

- Cameras with limited dynamic range

  Brightest pixel value / Darkest pixel value

Figure. High contrast scene.

- Dynamic range is the range of luminance supported by a medium.

$10^{-6}\ cd/m^2$ $10^{8}\ cd/m^2$

Human Visual System* (full range)

HVS* (simultaneous)

LDR

HDR

Background

Solutions for HDRI problem

HDR compatible devices

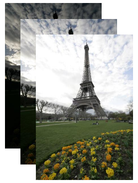

- Sensor production
- Not accessible for all due to economic cost

Tone-mapping [1]

LDR

- Problems arise in images [2, 3]
  - Low subjective contrast
  - Color saturation

MEF*



LDR

- Computationally efficient
- Preferred by user-grade device manufacturers

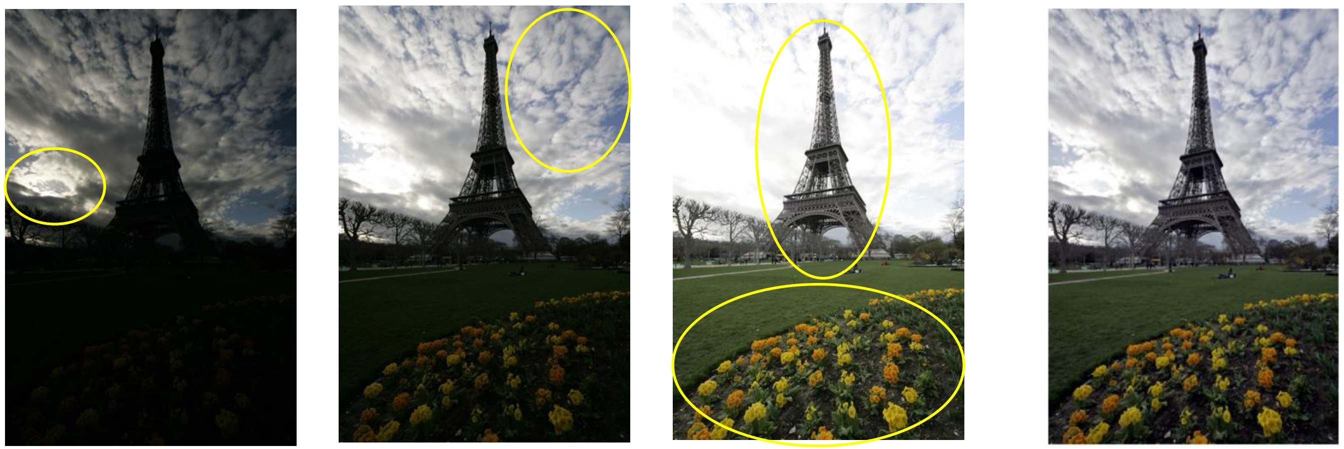
* Multi-exposure image fusion

1. Reinhard, E. *et al. Photographic tone reproduction for digital images*, in Conference on Computer Graphics and Interactive Techniques. San Antonio, Texas. July 2002.
2. Akyuz, A. O. and Reinhard, E. (2006) *Color appearance in high-dynamic-range imaging*, Journal of Electronic Imaging. International Society for Optics and Photonics, Vol.15(3), pp. 33001.
3. Kiser, C. *et al. Real time automated tone mapping system for HDR video*, in IEEE International Conference on Image Processing. Orlando, Florida. 30 September - 3 October 2012.

Background

Multi-exposure image fusion

- Combining a stack of input exposures of the same scene into a single informative HDR-like content [1,2].



Under-exposed   Normal-exposed   Over-exposed   Fused image

Figure. The main idea of MEF.

- Each exposure has distinct parts of details.
- Combining the exposures via weight maps.
  - ✓ Without damaging the fine details and color information.
- Acquired content can be projected to any LDR screen.

1. Mertens, T., Kautz, J. and Van Reeth, F. (2009) *Exposure fusion: a simple and practical alternative to high dynamic range photography*, Computer Graphics forum, Vol. 28(1), pp. 161–171.
2. Burt, P. J. and Kolczynski, R. J. *Enhanced image capture through fusion*, in International Conference on Computer Vision. Berlin, Germany. 11-14 May 1993.

Proposed Method

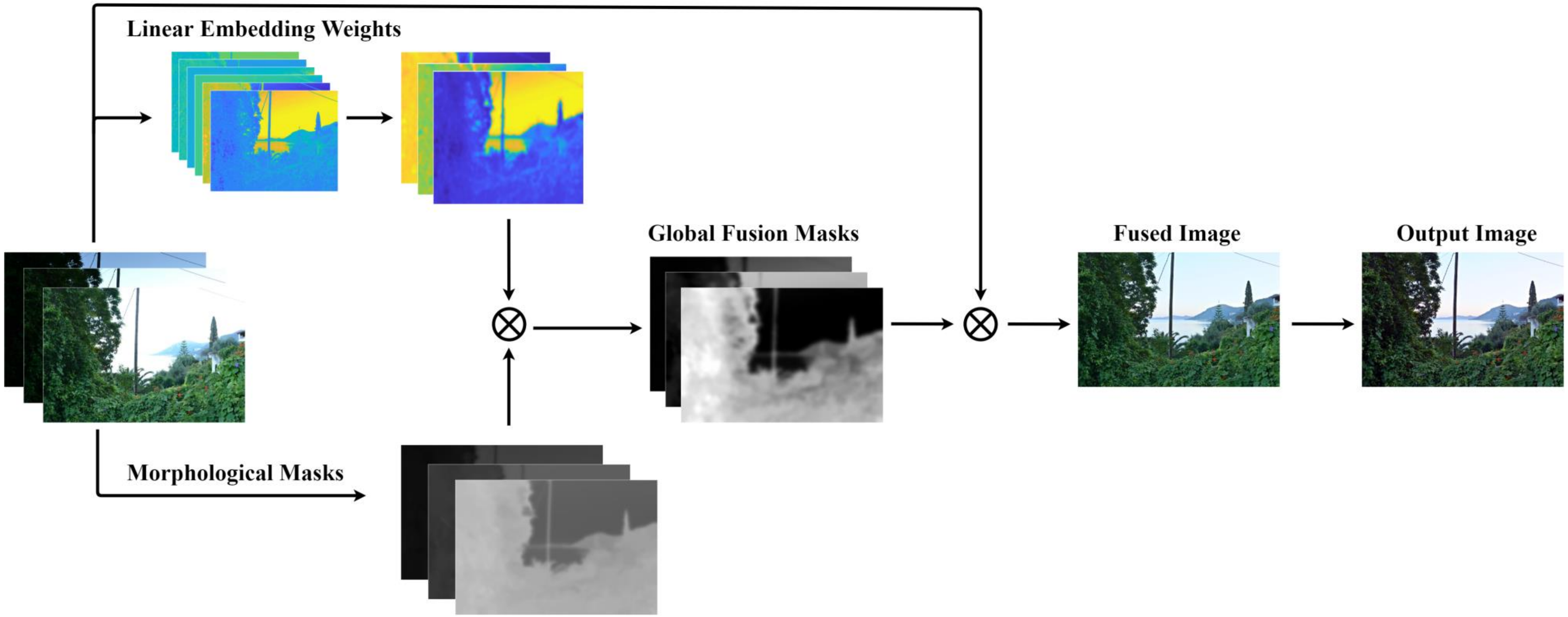


Figure. Basic flowchart of the proposed method.

Proposed Method

Determination of main exposures

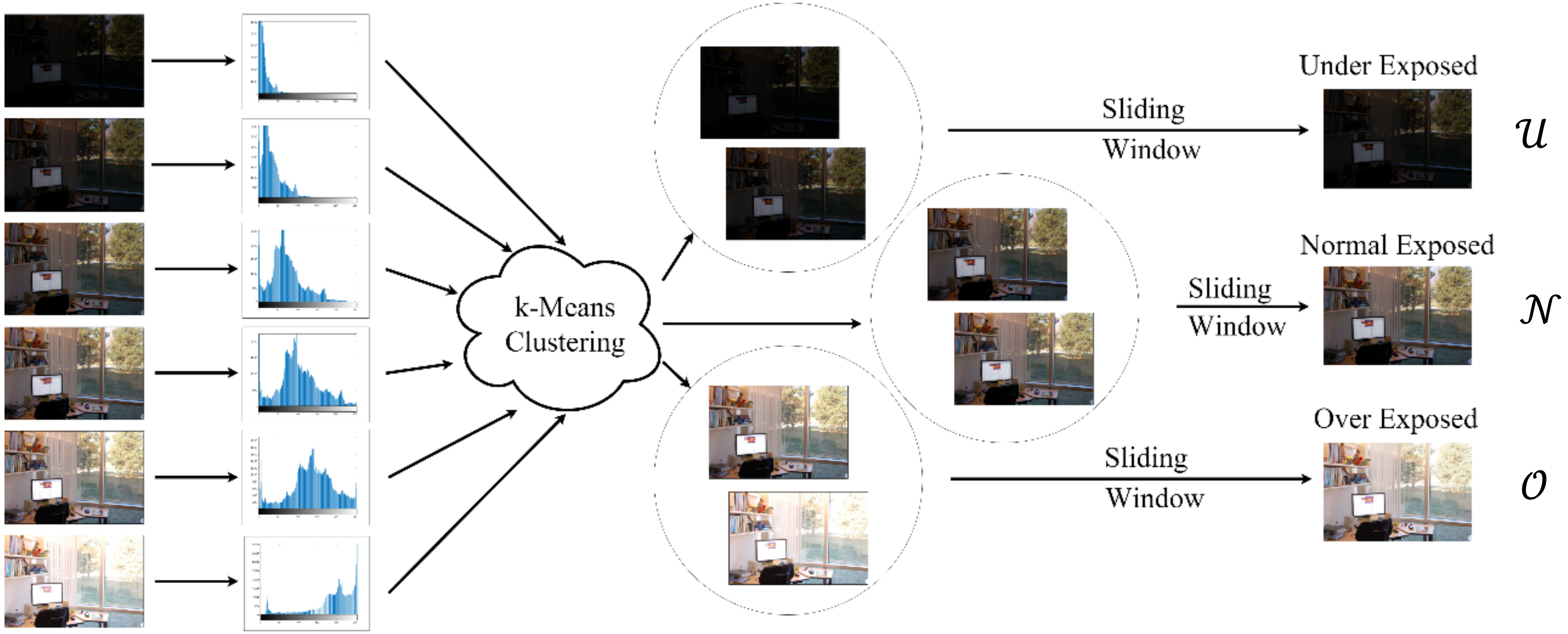


Figure. Scheme of the process.[1]

- Obtain PDFs to form feature vectors
- Apply k-means clustering method to group all exposures in the input stack to 3 main exposure sets
- Apply sliding window technique to exposure sets to obtain 3 exposures, under- and normal- and over-exposed

1) O.Ulucan , D.Karakaya and M.Turkan. (2021) Multi-exposure image fusion based on linear embeddings and watershed masking. *Signal Processing* , Vol. 178, No. 107791

Proposed Method
Weight maps via Linear embeddings

Linear Embeddings

- New framework for weight map characterization
- Inspired from LLE* [1]

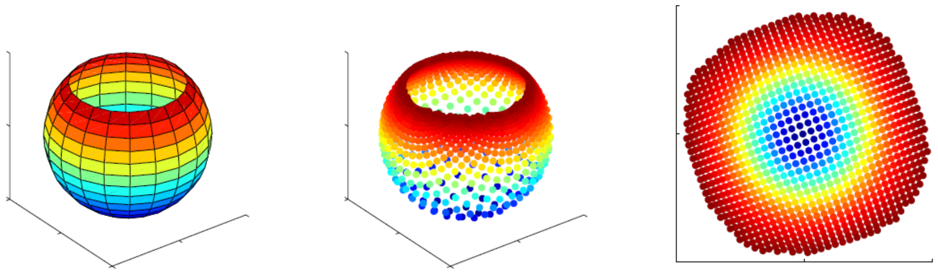


Figure. Neighbor preserving mapping via LLE [1] .

“Data points and its neighbors lie on or close to locally linear patch of the manifold” [1]

Each exposure is sampled from a manifold structure and all these exposures should lie on or close to a locally linear patch of the underlying sampled manifold

“Nearby points in the high dimensional space remain nearby and similarly co-located with respect to one another in the low dimensional space” [1]

LLE preserves the local geometry and structure of the manifold

* Locally Linear Embedding

1. Roweis, S. T. and Saul, L. K. (2000) *Nonlinear dimensionality reduction by locally linear embedding*, Science, Vol. 290(5500), pp. 2323–2326.

Proposed Method

Weight maps via Linear embeddings

 $= \mathbf{W}^1$  $+ \mathbf{W}^2$ 

$$\{\mathbf{W}_i^1, \mathbf{W}_i^2\} = \arg\min_{\{w_1,w_2\}} \left\| \boldsymbol{o}_i - [\boldsymbol{u}_i \;\; \boldsymbol{n}_i] \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2^2 \quad s.t \;\; w_1 + w_2 = 1$$

 $= \mathbf{W}^3$  $+ \mathbf{W}^4$ 

$$\{\mathbf{W}_i^3, \mathbf{W}_i^4\} = \arg\min_{\{w_3,w_4\}} \left\| \boldsymbol{u}_i - [\boldsymbol{n}_i \;\; \boldsymbol{o}_i] \begin{bmatrix} w_3 \\ w_4 \end{bmatrix} \right\|_2^2 \quad s.t \;\; w_3 + w_4 = 1$$

 $= \mathbf{W}^5$  $+ \mathbf{W}^6$ 

$$\{\mathbf{W}_i^5, \mathbf{W}_i^6\} = \arg\min_{\{w_5,w_6\}} \left\| \boldsymbol{n}_i - [\boldsymbol{u}_i \;\; \boldsymbol{o}_i] \begin{bmatrix} w_5 \\ w_6 \end{bmatrix} \right\|_2^2 \quad s.t \;\; w_5 + w_6 = 1$$

- To form linear embedding weight maps
- To maintain local smoothness in the transition regions while avoding possible noise and artifacts

$$\mathbf{E}_1' = (|\mathrm{W}^1| + |\mathrm{W}^5|) * \mathbf{G}$$

$$\mathbf{E}_2' = (|\mathrm{W}^2| + |\mathrm{W}^3|) * \mathbf{G}$$

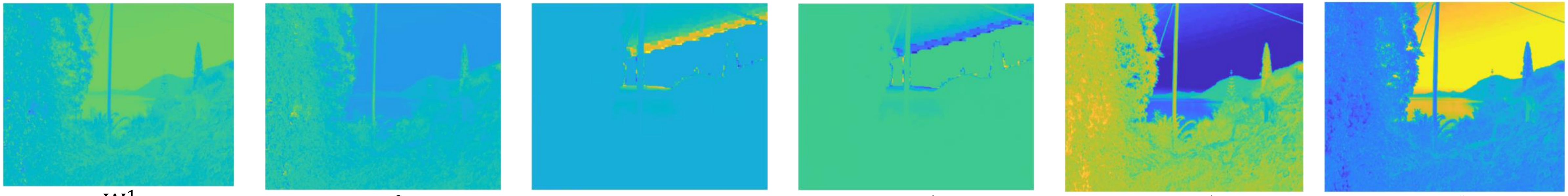
$$\mathbf{E}_3' = (|\mathrm{W}^4| + |\mathrm{W}^6|) * \mathbf{G}$$

$\mathbf{G}$: Gaussian smoothing kernel

$*$: Convolution operator

Proposed Method

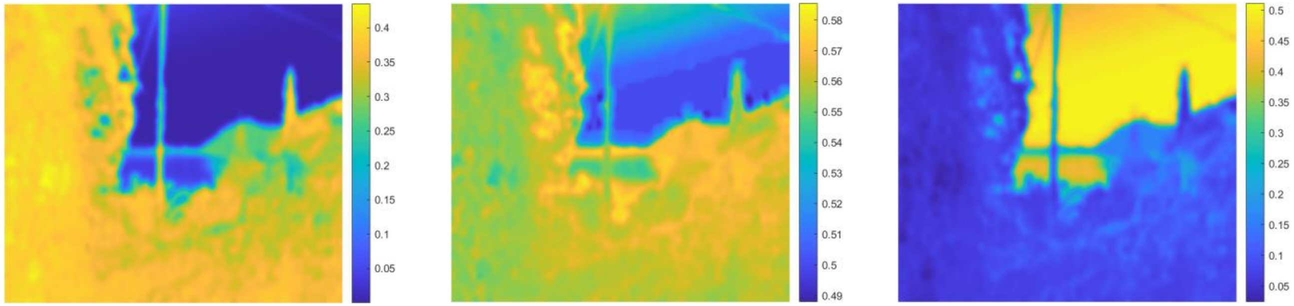
Weight maps via Linear embeddings



W^1 W^2 W^3 W^4 W^5 W^6

Figure. Optimal weight maps via LE.

$$\mathbf{E}_k = \mathbf{E}'_k \oslash (\mathbf{E}'_1 + \mathbf{E}'_2 + \mathbf{E}'_3), \ \ k = 1, 2, 3$$



$\mathbf{E}_1$ $\mathbf{E}_2$ $\mathbf{E}_3$

Figure. Weight maps via LE.

- Each weight map highlights specific parts of the exposures to be fused.

$\oslash$: Element-wise division

Proposed Method
Adaptive morphological masking

$$\text{Hat function} = \begin{cases} 1, & \beta < I < 255 - \beta \\ 0, & otherwise \end{cases}$$

- Artifacts in sharp texture and color changes

Smoothing filters

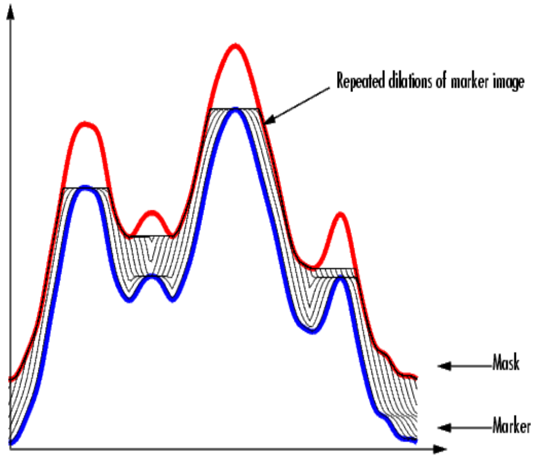
- Artifacts, halo effects

Edge-aware smoothing filters

- Hard to adjust parameters

Adaptive Morphological Masking

- Natural texture and color transitions
- Remove small blemishes without damaging overall structures



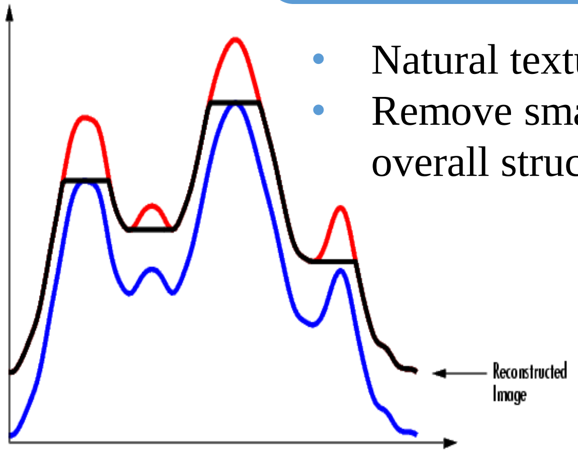


Figure. Example of halo effect.

Proposed Method

Adaptive morphological masking

\# of darkest pixel intensity of $\mathcal{U}$ < # of brightest pixel intensity of $\mathcal{O}$, r* = 20

\# of darkest pixel intensity of $\mathcal{U}$ > # of brightest pixel intensity of $\mathcal{O}$, r* = 11

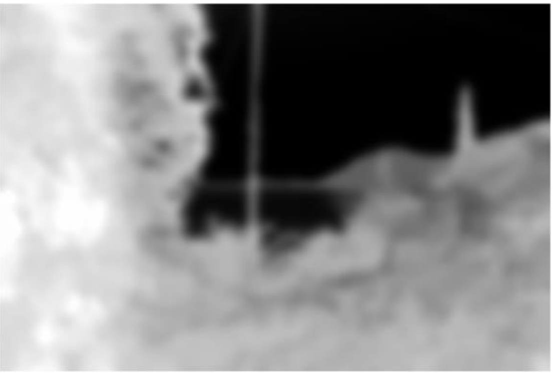
✓ Opening-by-reconstruction operation followed by closing-by-reconstuction is carried out

Figure. Morphological masks for *Flowers*: $\mathbf{M}_1$, $\mathbf{M}_2$, $\mathbf{M}_3$.

* Radius size of disk-shaped structuring element

Proposed Method

Exposure fusion and post processing



$\mathbf{G}_1 = \mathbf{M}_1 \otimes \mathbf{E}_3$

$\mathbf{G}_2 = \mathbf{M}_2 \otimes \mathbf{E}_2$

$\mathbf{G}_3 = \mathbf{M}_3 \otimes \mathbf{E}_1$

- ✓ Global fusion masks are formed by exchanging the linear embedding weight maps of $\mathcal{U}$ and $\mathcal{O}$ via the morphological masks in order to highlight well-exposed areas in $\mathcal{U}$ and $\mathcal{O}$

- ✓ the top 1% and the bottom 1% of all pixel values of $\mathbf{G}_2$ are clipped to stretch the contrast and obtain a more balanced contribution from $\mathcal{N}$

$$\mathbf{F} = \mathrm{U} \otimes \mathbf{G}_1 + \mathrm{N} \otimes \mathbf{G}_2 + \mathrm{O} \otimes \mathbf{G}_3$$

- ✓ the top 1% and the bottom 1% of all pixel values of $\mathbf{F}$ are saturated to recover small low-light areas and mediocre color intensities

$\otimes$: Element-wise multiplication

Experimental Results

Dataset

Ma, Duanmu, Yeganeh and Wang [1]
- 10 different static image stacks
  - Distinct number of exposures
  - Low resolution

Merianos and Mitianoudis [2]
- 2 different static image stacks
  - 3 exposures
  - High resolution

Erdem Okur
- 2 different static image stacks
  - 3 exposures
  - Low Resolution
- New images will be avaliable for your research @ Github*

Table. The dataset used in this study.

Name	*Arno*	*Chinese Garden*	*Church*	*Farmhouse*	*Flowers*	*Landscape*	*Laurenziana*
Size	$339 \times 512 \times 3$	$340 \times 512 \times 3$	$512 \times 335 \times 3$	$340 \times 512 \times 3$	$720 \times 1080 \times 3$	$341 \times 512 \times 3$	$512 \times 356 \times 3$
Name	*Mask*	*IzmirNight*	*Office*	*OldHouse*	*IzmirFair*	*Tower*	*Venice*
Size	$341 \times 512 \times 3$	$518 \times 690 \times 3$	$340 \times 512 \times 6$	$720 \times 1080 \times 3$	$456 \times 342 \times 3$	$512 \times 341 \times 3$	$341 \times 512 \times 3$

* https://github.com/DiclehanOguzhan

1. K. Ma, Z. Duanmu, H.Yeganeh, and Z.Wang, "Multi-exposure image fusion by optimizing a structural similarity index," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 60–72, December 2017.
2. I. Merianos and N. Mitianoudis, "Multiple-exposure image fusion for HDR image synthesis using learned analysis transformations," *J. Imaging*, vol. 5, no. 3, pp. 32, February 2019.

Experimental Results

✓ Proposed algorithm compared with Mertens [1], Paul [2], Ma [3], Li18 [4], Lee [5], Liu [6], Hayat [7], Li20 [8]

- AMD Ryzen(TM) 5 3600x CPU @ 3.80GHz 6-core
- 16GB RAM
- MATLAB R2019b

MEF-SSIM*

- Measures patch structural consistency → Contrast and Structure
- Luminance in fused image is considered

0  1

* K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3345–3356, June 2015.
1. T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion: a simple and practical alternative to high dynamic range photography," *Comp. Graph. Forum*, vol. 28, no. 1, pp. 161–171, February 2009.
2. S. Paul, I. S. Sevcenco, and P. Agathoklis, "Multi-exposure and multi-focus image fusion in gradient domain," *J. Circuit Syst. Comp.*, vol. 25, no. 10, pp. 1650123, June 2016.
3. K. Ma, H. Li, H. Yong, Z.Wang, D. Meng, and L. Zhang, "Robust multi-exposure image fusion: a structural patch decomposition approach," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2519–2532, February 2017.
4. H. Li and L. Zhang, "Multi-exposure fusion with CNN features," in *IEEE Int. Conf. Image Process.*, 2018, pp. 1723–1727.
5. S. Lee, J. S. Park, and N. I. Cho, "A multi-exposure image fusion based on the adaptive weights reflecting the relative pixel intensity and global gradient," in *IEEE Int. Conf. Image Process.*, 2018, pp. 1737–1741.
6. Q. Liu and H. Leung, "Variable augmented neural network for decolorization and multi-exposure fusion," *Inf. Fusion*, vol. 46, no. 1, pp. 114–127, March 2019.
7. N. Hayat and M. Imran, "Ghost-free multi exposure image fusion technique using dense SIFT descriptor and guided filter," *J. Vis. Commun. Image Represent.*, vol. 62, pp. 295–308, July 2019.
8. H. Li, K. Ma, H. Yong, and L. Zhang, "Fast multi-scale structural patch decomposition for multi-exposure image fusion," *IEEE Trans. Image Process.*, vol. 29, pp. 5805–5816, April 2020.

Experimental Results

Table. MEF-SSIM scores.

	Algorithms								
	Mertens	Paul	Ma	Li18	Lee	Liu	Hayat	Li20	Proposed
Arno	**0.991**	0.958	0.980	0.948	0.987	0.985	0.985	0.990	0.986
Chinese Garden	0.989	0.982	0.985	0.977	0.990	0.988	0.993	**0.994**	0.991
Church	0.989	0.978	**0.992**	0.980	**0.992**	0.977	**0.992**	**0.992**	0.991
Farmhouse	0.981	0.971	0.984	0.984	0.979	0.978	0.984	**0.986**	0.983
Flowers	0.964	0.961	0.987	0.972	0.990	0.990	**0.995**	**0.995**	0.991
Landscape	0.976	0.972	0.993	0.954	0.981	**0.994**	0.973	0.988	0.986
Laurenziana	0.988	0.982	0.985	0.973	0.987	0.987	0.989	**0.990**	0.989
Mask	0.987	0.975	0.988	0.975	0.990	0.985	**0.992**	**0.992**	0.987
IzmirNight	0.952	0.984	0.989	0.964	0.988	0.988	0.989	0.991	**0.992**
Office	0.985	0.973	0.988	0.970	**0.991**	0.985	0.987	0.990	**0.991**
OldHouse	0.974	0.973	0.987	0.962	0.990	0.988	0.968	0.990	**0.991**
IzmirFair	0.950	0.983	0.992	0.976	0.990	0.992	0.993	**0.996**	0.992
Tower	0.986	0.977	0.986	0.981	0.987	0.983	0.987	**0.988**	0.984
Venice	0.966	0.954	0.940	0.947	0.972	0.973	0.972	**0.984**	0.979
avg	0.977	0.973	0.984	0.967	0.987	0.985	0.986	**0.990**	0.988
std	0.0139	0.0094	0.0131	0.0120	0.0056	0.0058	0.0086	**0.0033**	0.0040

Experimental Results

Figure. Visual comparison of different methods for *Venice*. (Left-to-right) Input stack; Liu (0.973); Li20 (0.984); Proposed (0.979).

Experimental Results

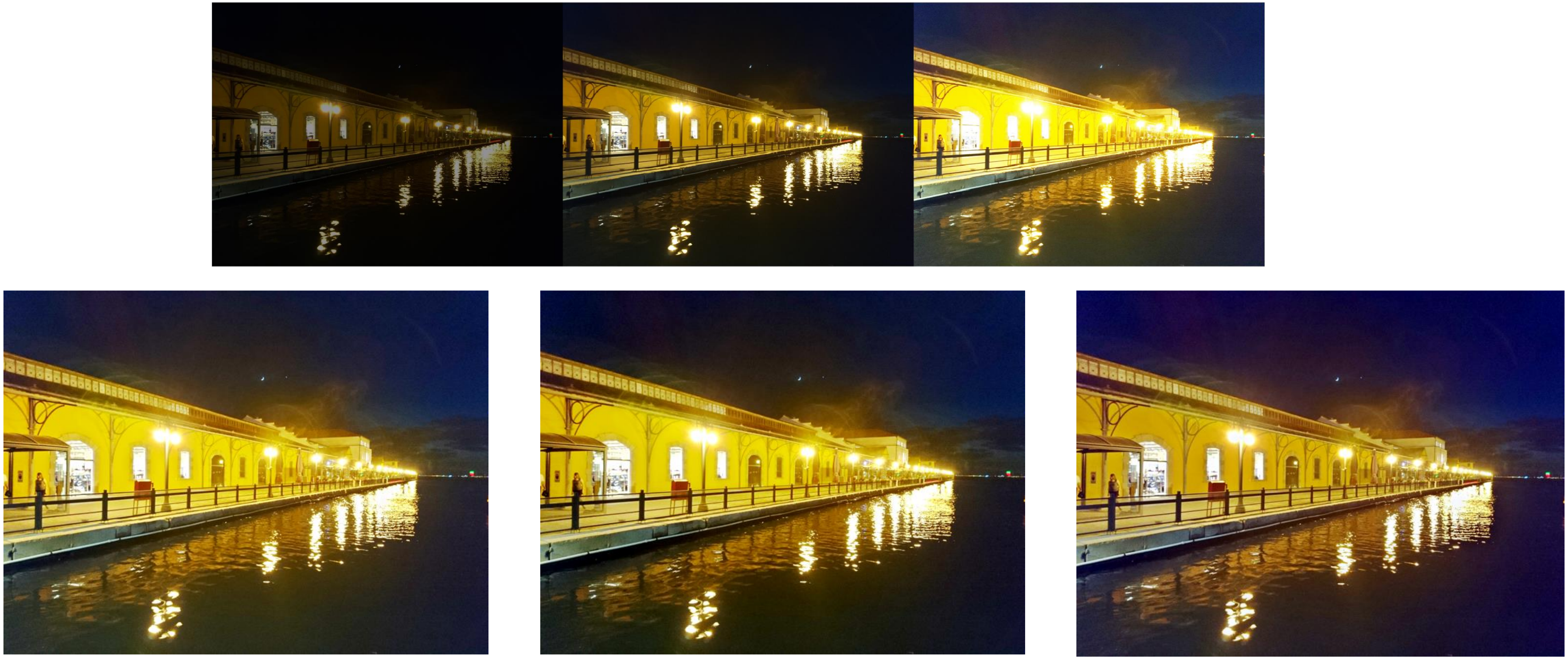


Figure. Visual comparison of different methods for *IzmirNight*. (Left-to-right) Input stack; Ma (0.989); Li20 (0.991); Proposed (0.992).

Experimental Results
Application to visible and infrared image fusion

- ✓ Visually compared with Liu18 [1], Li19 [2], Bavirisetti [3]
- ✓ Proposed algorithm not modified for this application

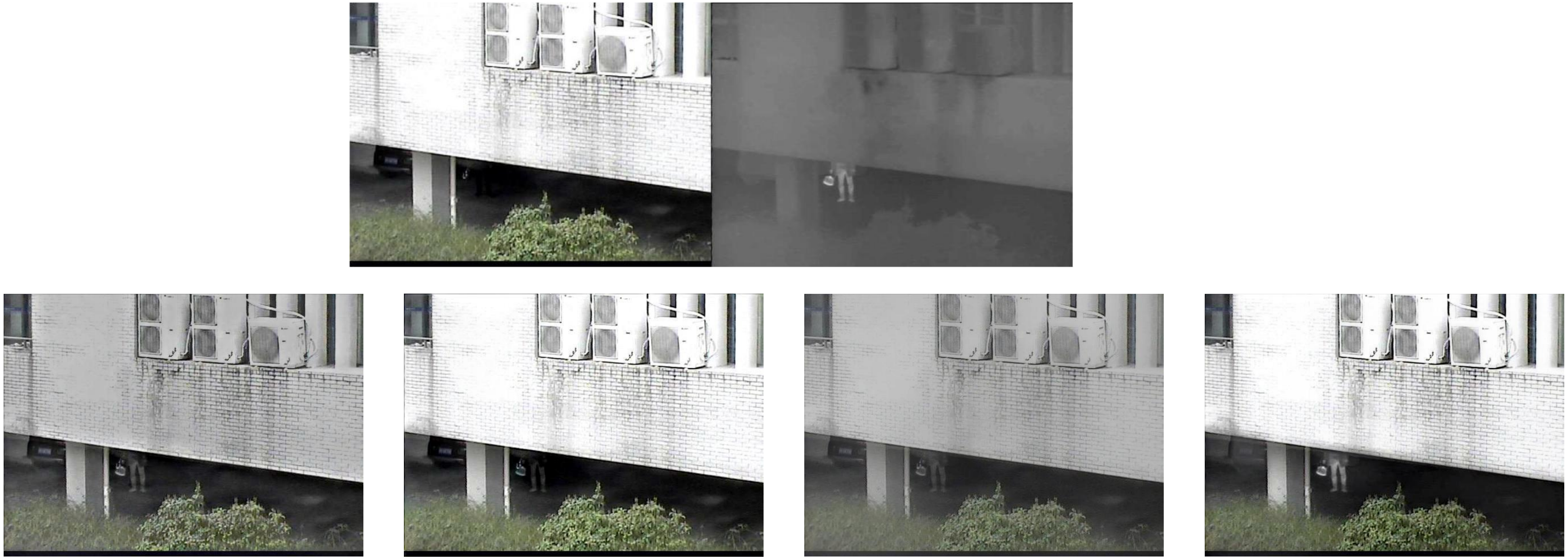


Figure. Visual comparison of different methods for *Kettle*. (Top) The input pair. (Bottom) Bavirisetti; Liu18; Li19; Proposed.

Conclusion

- ✓ New weight map characterization framework proposed
  - ➢ Linear embeddings
  - ➢ Morphological masking
- ✓ Degined method presents highly competitive visual results
- ✓ Proposed algorithm has not been modified for visible and infrared image fusion, but proposed weight map characterization process has shown its potential in this domain
- ✓ Algorithm will be further investigated for other image fusion applications
- ✓ New images, and the outcomes with the code of proposed method can be reached @ Github*

* https://github.com/DiclehanOguzhan

Thank you for listening!

Questions?