

Efficient High Dynamic Range Imaging via Matrix Completion

Grigorios Tsagkatakis*, and Panagiotis Tsakalides*[†]

*Institute of Computer Science (ICS) - Foundation for Research & Technology - Hellas (FORTH), Crete, Greece

[†]Department of Computer Science, University of Crete, Heraklion, Crete, Greece

e-mail: {greg, tsakalid}@ics.forth.gr

Abstract—Typical digital cameras exhibit a limitation regarding the dynamic range of the scene radiance they can capture. High Dynamic Range (HDR) imaging refers to methods and systems that aim to generate images that exhibit higher dynamic range between the lightest and the darkest parts of an image. A typical approach for generating HDR images is exposure bracketing where multiple frames, each one with a different exposure setting, are captured and combined to a HDR image of the scene. The large number of images that exposure bracketing requires often leads to motion artefacts that limit the visual quality of the resulting HDR image. In this work, we propose a novel approach in HDR imaging that significantly reduces the necessary number of images.

In our proposed system, each pixel of a single frame collects light for a random amount of time i.e. random exposure. By collecting a small number of such images, the full sequence of low dynamic range images can be reconstructed and subsequently used for HDR generation. The problem is solved by casting the reconstruction of the sequence as a nuclear norm minimization problem following the premises of the recently proposed theory of Matrix Completion. Experimental results suggest that the proposed method is able to reconstruct the sequence from as low as 20% of the images that traditional techniques require with minimal reduction in image quality.

Index Terms—High Dynamic Range Imaging, Low rank, Matrix Completion

I. INTRODUCTION

Imaging sensors found in digital cameras exhibit limitations regarding the range of light intensity they can capture. Dynamic range refers to the range between the minimum and the maximum luminance that is captured or emitted by a device. For a camera, the dynamic range refers to the luminance range between minimum noise level and saturation level. Depending on the configuration of the camera, regions of the captured image may be under-exposed, i.e. present noise-like characteristics, while other regions may be over-exposed, i.e. information is lost due to saturation. The under- or over-exposure of an image can lead to texture losses, color quality reduction and an overall decrease in perceptual image quality. High Dynamic Range imaging (HDR) refers to methods that employ specialized techniques in order to enhance the dynamic range between the lightest and the darkest part of an image and produce higher quality images. In addition to the artistic value of HDR imaging, HDR imaging has been successfully

employed in numerous applications including physically-based rendering, remote sensing and specialized imaging including biological, astronomical and palaeontological imaging.

In general, HDR imaging can be performed using specialized hardware that directly acquire HDR images or employ a specialized technique to extract HDR information from a single or multiple low dynamic range images. A well known technique for HDR imaging that can be applied using off-the-shelf traditional Low Dynamic Range (LDR) cameras is *exposure bracketing*. In exposure bracketing, a large number of images of the same scene is captured, each one with different exposure characteristics. Figure 1 presents an example of images taken with different exposure setting.



Figure 1. Images of the same static scene under three exposure settings.

Under assumptions such as stationary of the scene and the camera, these multiple low dynamic range images can be combined in order to produce the HDR image. For cases where the displaying device is not able to fully reproduce the HDR image, a *tone mapping* process is applied where the optimum exposure for each part of the scene is combined to generate an LDR image with HDR-like characteristics.

As one can expect, a major limitation of exposure bracketing for HDR imaging is the scene stationary requirement. In this work, we propose a novel approach in HDR acquisition that exploits the spatio-temporal redundancies that exist between the differently exposed images in a principled way in order to dramatically reduce the number of necessary image acquisitions. Our proposed system assumes that the camera system is able to control the exposure of each individual pixel element either explicitly or implicitly by employing an appropriate filter.

The rest of the paper is organized as follows. In Section II, we briefly describe the state-of-the-art in HDR imaging and discuss the connections between these methods and the proposed one. We provide a motivation for this work in Section III. Section IV outlines the theoretical framework of matrix completion while Section V discusses the application of the MC framework in the HDR imaging problem. Experimental

results are presented in Section VI and the paper is concluded in Section VII.

II. RELATED WORK

Following the discussion in the previous section, a major limitation of exposure bracketing techniques is the requirement for capturing multiple image. [7] was one of the first thorough investigations of minimal bracketing sets for HDR imaging. In that work, the authors investigated three scenarios of exposure control based on the amount of available scene information and the availability of real-time processing. Based on these scenarios they provided optimality conditions for two types of exposure control, namely geometric exposure variation and arithmetic exposure variation.

In [5], the authors investigated the optimal exposure of the capture sequence in order to achieve an optimal worst-case performance. The problem of optimal parameter settings was formulated as an integer programming optimization according to a detailed noise model that encoded photon noise as well as pre and post ISO gain noise. An extreme case of minimum exposure bracketing is the case of single shot HDR imaging [6] where optimally exposed regions from each color channel were fused in order to generate the HDR image. Although this method did not achieve the same quality as time multiplexed methods, it does not require specialized hardware nor does it suffer from motion artefacts.

While the previous methods address the generation of HDR methods, in many scenarios, the images are mapped back to a low dynamic range, limited by the display capabilities. For this class of scenarios, the process could bypass the generation of the HDR image and process directly to a higher quality LDR image. Such a design was recently presented in [4] where fusion of the multiple exposures was used for generating the high quality image.

A concept for obtaining HDR images, similar to the one presented in this paper, was proposed in 2000 by Nayer and Mitsunaga [18] termed Spatially Varying Exposure (SVE). In SVE a single image acquisition is performed. However, in the captured image, each pixel contains multiple levels of exposure by placing a mask with different optical transparencies adjacent to the sensor. Reconstruction of the HDR images is achieved by a cubic interpolation method. The proposed methods is significantly different both in terms of acquisition where one value is taken for each pixel and reconstruction where the novel theory of Matrix Completion is employed.

III. MOTIVATION

Let E be the scene radiance reaching the camera sensor after going through the system's optics. Given the exposure time T , the sensor exposure captured by the sensor is modelled as $I = E * T$. The photons reaching the sensor for a given amount of exposure will undergo further processing, such as amplification and analog-to-digital conversion, and will be eventually translated into a small, typically 8 or 12 bits, brightness value B . This conversion process is modelled by a function called camera response function (CRF). Let the

CRF function be denoted as \mathcal{F} i.e. $B = \mathcal{F}(I)$. Identifying the CRF function \mathcal{F} , a process termed radiometric calibration, offers many benefits including radiometric corrections, photometric stereo and shape-from-shading. Developing realistic CRF models has been extensively explored in the literature and various approaches have been considered.

Mann et al. [15] were one of the firsts to explore modelling of CRF. They proposed a simple gamma function for approximating the CRF. Debevec et al. [16] extended the previous work by considering a more detailed model where a set of exposure values were utilized in solving a regularized least squares problem and providing a description of the CRF curves. Mitsunaga et al [17] proposed a low order polynomial for expressing the CRF curve. More recently, in [8], the authors explored the space where CRF functions reside in order to obtain a robust mapping between scene radiance and image intensity and facilitate CRF interpolation. Utilizing information extracted from a collection of 201 real-world response functions, the authors suggested that CRFs can be modelled by low order polynomials.

This work builds on the premisses of previous work and suggests a computationally efficient method for reconstructing the full collection of LDR images from a subset of space-exposure multiplexed set of images. Let $\mathbf{P} = \{P_1^0, P_1^1, \dots, P_1^s, \dots, P_k^s\}$ be a spatio-temporal group of neighbouring pixels where the superscript indexing refers to spatial location and the subscript indexing refers to the time/exposure. By vectorizing the spatio-temporal group along the temporal direction, a matrix $M \in \mathbb{R}^{k \times s}$ can be formed as $M = [vec(P^1) \dots vec(P^k)]$, where each column $vec(P)$ corresponds to a specific exposure. The matrix M can be rewritten as $M = [vec(\mathcal{F}(I^1)) \dots vec(\mathcal{F}(I^k))] = [vec(\mathcal{F}(I * E^1)) \dots vec(\mathcal{F}(I * E^k))]$, where in this case I corresponds to the irradiance of the group of pixels.

Based on the above formulation, the rank of the matrix M will depend on the characteristics of the CRF function \mathcal{F} . If \mathcal{F} is a linear function, then a decomposition $\mathcal{F}(I * E^k) = \mathcal{F}(I) * \mathcal{F}(E^k) = F(I) * w$ is possible where w is a scalar number proportional to the exposure. Assuming the exposure times are known, the rank of matrix M is one, since all columns can be generated by the multiplication of the first column with a scalar w corresponding to the exposure. In reality, the CRF cannot be adequately modelled by a linear function but instead requires a non-linear model. Examples of this case of models include the low order polynomial suggested in [17] or the model proposed in [8]. A common theme of the papers is that the CRF's intrinsic dimensionality is much lower compared to the explicit dimensionality of the data. Furthermore, the spatial correlation that exists within a pixel neighbourhood also contributes to the a lower dimensionality. As a result of the spatio-temporal correlations, the data matrix exhibits a small degree of freedom that is manifested by the reduced rank of the matrix.

To further support the argument that the matrix M is a low rank matrix, we collected patches of size 5×5 pixels from a large number of LDR sequences, captured at 28 different

exposure setting. To identify the rank of the matrix, Singular Value Decomposition was applied to each patch and Figure 2 presents the average singular values of the patches along with an error bar encoding the standard deviation. As one can observe, only a small number of singular values are active i.e. most of the variance/energy of the signal is captured by the first 4 – 5 singular values indicating a low rank matrix.

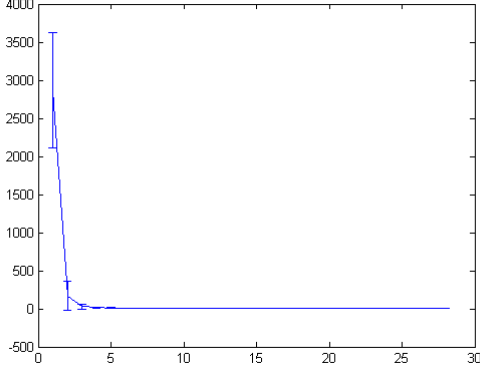


Figure 2. Singular Values.

IV. LOW RANK MATRIX RECOVERY

Let $\mathbf{X} = [X_1, X_2, \dots, X_s] \in \mathbb{R}^{n \times s}$ be a measurement matrix consisting of s measurements from n different sources. In general, one cannot recover the ns entries of the matrix \mathbf{X} from a smaller number of k entries, where $k \ll ns$. However, it has been recently shown that such a recovery is possible when the rank of the matrix is small compared to the dimensions of the matrix. More interestingly, one can recover the matrix \mathbf{X} from $k \geq Cn^{6/5}r \log(n)$ random measurements, where $n > s$ and $\text{rank}(\mathbf{X}) = r$ [1]. To recover the unknown matrix X from the measurements M , the following minimization problem needs to be solved

$$\min\{\text{rank}(\mathbf{M}) : \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{M})\} \quad (1)$$

where \mathcal{A} is a linear map from $\mathbb{R}^{ns} \rightarrow \mathbb{R}^k$. The theory of Matrix Completion (MC) suggests that recovery is possible when the linear map \mathcal{A} is defined as a random sampling operator that records a small number of entries from the matrix

$$\mathcal{A}_{ij}(\mathbf{X}) = \begin{cases} X_{ij} & \text{if } ij \in S \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where S is the set of sampled values. In general, to solve the MC problem, the linear map \mathcal{A} must satisfy a modified Restricted Isometry Property, which is the case for uniform Random Sampling in both rows and columns [2].

Although solving the above problem will generate a low rank matrix consistent with the observations, rank minimization is an NP-hard problem and therefore cannot be used in practice. Recently, a relaxation of the above problem [1] was shown to produce arbitrary accurate results, by replacing the rank constraint by the tractable nuclear norm minimization

$$\min\{\|\mathbf{M}\|_* : \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{M})\} \quad (3)$$

where the nuclear norm is defined as $\|\mathbf{M}\|_* = \sum \|\lambda_i\|_1$ i.e. the ℓ_1 norm of the singular values. For the noisy case, an approximate version can be solved [3]

$$\min\{\|\mathbf{M}\|_* : \|\mathcal{A}(\mathbf{X}) - \mathcal{A}(\mathbf{M})\|_F^2 \leq \epsilon\} \quad (4)$$

where $\|\mathbf{X}\|_F^2 = \sum \lambda_i^2$ denotes the Frobenius norm and ϵ is the acceptable approximation error.

To solve the nuclear norm minimization problem in (3) and (4), various approaches have been proposed including the SVT [9], the ALM [10] and the OptSpace [11]. Furthermore, the application of MC on various imaging problems has been recently explored including problems such as video denoising [12], image segmentation [13] and multi-label image classification [14] among others.

V. MATRIX COMPLETION FOR HDR IMAGING

To reduce the number of necessary LDR images, we propose the application of a random sampling scheme, where in each frame, each pixel randomly selects the exposure time. As a result, for each pixel, only a small number of exposure values are acquired, while neighbouring pixels record different subset of exposures.

Formally, Let $\mathbf{I} = \{\mathbf{I}^1, \mathbf{I}^2 \dots \mathbf{I}^T\}$ be a set of LDR images where each image \mathbf{I}^t corresponds to a different exposure setting. Furthermore, each image $\mathbf{I}^t = \{B_1, B_2, \dots B_K\}$ can be divided into K , possibly overlapping, spatial blocks. In our model we assume that for each block, only a small number of measurements are available. We formulate the problem of recovering all the measurements as a nuclear norm minimization problem:

$$\min\{\|\mathbf{X}\|_* : \|\mathcal{A}(\mathbf{X}) - \mathcal{A}(\mathbf{B}_i^j)\|_F^2 \leq \epsilon\} \quad (5)$$

Figure 3 presents a block diagram of the propose system. We note that block sizes are overgrown for exposition purposes.

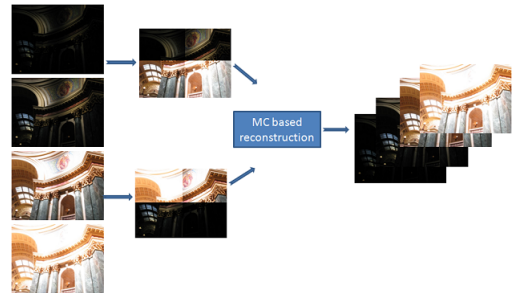


Figure 3. Block diagram of the proposed system.

To further reduce artefacts that can be produced during reconstruction, the final pixel value in each exposure is generated by averaging the corresponding value from the overlapping windows. Formally, the value of a pixel \hat{x}_{ijt} at spatial location ij and exposure setting t is found by $\hat{x}_{ijt} = \frac{1}{N} \sum_{i=1}^N x_{ij}$, where N is the number of overlapping windows at this location.

At this point, we note some of the benefits of the proposed approach:

- By reducing the number of necessary exposures, the proposed technique reduces the effects of motion such as motion artefacts and ghosting (appearance of the same object at multiple locations).
- Furthermore, the proposed system does not require significant hardware interventions. Current imaging system could be enhance by placing a filter or a Micro Mirror Device (MMD) .
- In addition, the proposed methods does not require the explicit estimation or definition of a camera response function that could be unavailable but performs an implicit estimation via the low rank assumption.

VI. EXPERIMENTAL RESULTS

To validate the proposed approach, a series of experiments was conducted where we utilized the LDR sequence of C. S. Verma and Mon-Ju . We only considered the luminance component but extending the process to coloured and multispectral images can be achieved by processing each color/frequency channel individually and combining the results.

We examined sampling rates between 20 – 50% corresponding to capturing 6 – 14 frames correspondingly, instead of the sequence of 28 image in the dataset. Table I presents numerical results on the reconstruction quality, measured in PSNR db, for three image sets and different sampling rates. These results suggest that the proposed method is able to achieve high quality and stable reconstruction from only a small set of images while increasing the captured frames offers little additional quality improvements.

sampling rate	0.2	0.3	0.4	0.5
set 1	24.3	27.0	27.5	27.7
set 2	25.2	29.5	31.1	31.6
set 3	24.6	27.8	28.5	28.7

Table I

PSNR VALUES FOR SAMPLING RATES 20% - 50 % ON THREE IMAGE SETS

Figures 4 and 5 present two LDR reconstructed examples for sampling rates 20 – 50% and different exposure setting. From these examples we can make a few observation. First, we observe that in short exposure times, corresponding to the top rows, the recovered images exhibit some blurring. This blurring is attributed to the averaging the takes place during the reconstruction. This effect could be reduced by performing a more intelligence pixel value estimation. In addition, we observe that in long exposures, bottom rows, the pixel values do not suffer the saturation that the original images exhibit. Last, regarding the sampling rate, we observe that a salt-and-pepper like noise is present in very low sampling rates (20%) but this noise is not present in higher sampling rates starting from 30%.

VII. DISCUSSION

Traditional method for HDR imaging require the acquisition of multiple LDR frames at different exposures in order to generate the HDR image. In this paper we propose a method

for capturing a sequence of LDR image that requires significantly less frames compared to traditional methods. In the proposed method, each pixel's exposure time is randomly selected, resulting in images containing pixels from various exposures. Using a small number of such images, we employ a recently proposed method, termed Matrix Completion, for estimating the missing exposure values and reconstructing the LDR sequence. The application of MC in this problem is motivated by the low rank nature of the space-time (exposure) cubes and employs a nuclear norm minimization. Experimental results suggest that reconstruction with minimum error is possible from 20% of the measurements that traditional methods require.

REFERENCES

- [1] E.J. Candès, and B. Recht, "Exact matrix completion via convex optimization", in *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717-772, 2009.
- [2] B. Recht, M. Fazel, and P.A. Parrilo, "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization", *SIAM Review*, vol. 52, No. 471, 2010.
- [3] R. Mazumder, T. Hastie, and R. Tibshirani, "Spectral regularization algorithms for learning large incomplete matrices", in *The Journal of Machine Learning Research*, vol. 11, pp. 2287-2322, 2010.
- [4] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion", in *IEEE Pacific Conference on Computer Graphics and Applications*, pp. 382-390, 2007.
- [5] S.W. Hasinoff, F. Durand, and W.T. Freeman, "Noise-optimal capture for high dynamic range photography", in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 553-560, 2010.
- [6] K. Hirakawa and P.M. Simon, "Single-shot high dynamic range imaging with conventional camera hardware", in *IEEE International Conference on Computer Vision (ICCV)*, pp. 1339-1346, 2011.
- [7] N. Barakat, A.N. Hone, and T.E. Darcie, "Minimal-bracketing sets for high-dynamic-range image capture", in *IEEE Transactions on Image Processing*, vol. 17(10), pp. 1864-1875, 2008.
- [8] Grossberg, M.D. and Nayar, S.K., "What is the space of camera response functions?", in *IEEE Conference on Computer Vision and Pattern Recognition*, vol. (2), pp. 602-609, 2003.
- [9] Cai, J.F. and Candès, E.J. and Shen, Z., "A Singular Value Thresholding Algorithm for Matrix Completion", in *SIAM Journal on Optimization*, vol. 20, pp. 1956-1982, 2010.
- [10] Lin Z., Chen M., Wu L., and Ma Y. "The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices", in *UIUC Technical Report UIUC-ENG-09-2215*, November 2009.
- [11] Keshavan, R.H. and Montanari, A. and Oh, S., "Matrix completion from noisy entries", in *The Journal of Machine Learning Research*, vol. 11, pp. 2057-2078, 2010
- [12] Ji, H. and Liu, C. and Shen, Z. and Xu, Y., "Robust video denoising using low rank matrix completion", in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1791-1798, 2010.
- [13] Cheng, B. and Liu, G. and Wang, J. and Huang, Z. and Yan, S., "Multi-task low-rank affinity pursuit for image segmentation", in *2011 IEEE International Conference on Computer Vision* pp. 2439-2446, 2011.
- [14] R.S. Cabral, F. De la Torre, J.P. Costeira and A. Bernardino, "Matrix Completion for Multi-label Image Classification", in *Advances in Neural Information Processing Systems (NIPS)*, 2011.
- [15] S. Mann and R. Picard, "On being undigital with digital cameras Extending dynamic range by combining differently exposed pictures", in *IS&T Society for Imaging Science and Technology Annual Conference*, vol. 56, pp. 24462461, May 1995.
- [16] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs", in *Proc. of SIGGRAPH 97, Computer Graphics Proc.*, pp. 369378, 1997.
- [17] T. Mitsunaga and S. K. Nayer, "Radiometric self calibration", in *IEEE Conf. on Computer Vision and Pattern Recognition*, vol.1, pp. 374380, Jun. 1999.
- [18] Nayar, S.K. and Mitsunaga, T. "High dynamic range imaging: Spatially varying pixel exposures", in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 472-479, 2000.

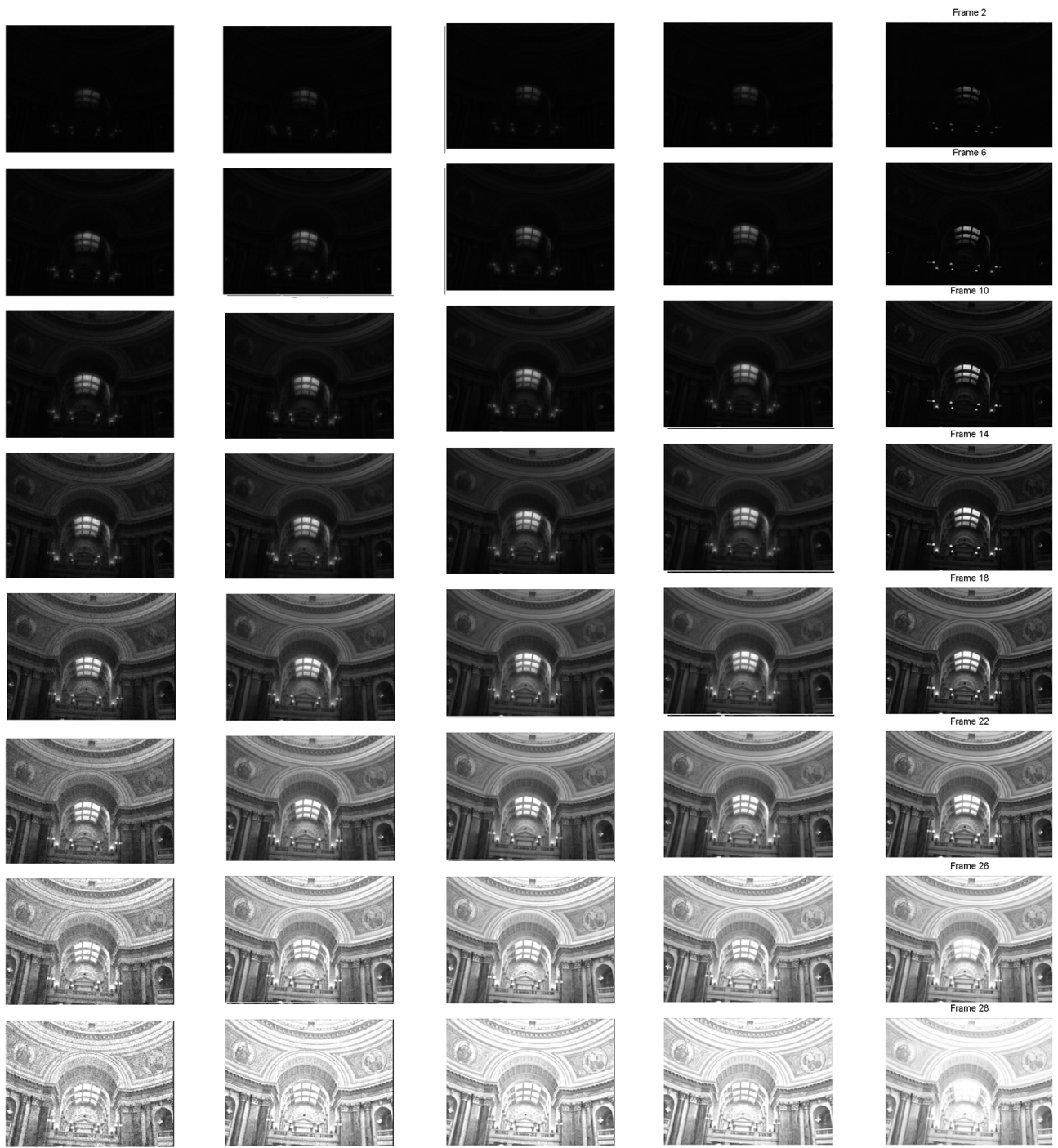


Figure 4. Recovered images. Columns correspond to sampling rates 20%, 30%, 40%, 50%, and original. Rows correspond to different exposure times.



Figure 5. Recovered images. Columns correspond to sampling rates 20%, 30%, 40%, 50%, and original. Rows correspond to different exposure times.