Introduction

- CMOS sensors are prevalent nowadays, especially in mobile phones, due to their lower cost and power consumption
- Sequential exposure of rows of sensors in CMOS cameras leads to rolling shutter (RS) effect
- Super-resolution (SR) from such images is a challenging task
- First attempt for the task of SR in CMOS cameras
- An RS-SR observation model that explains the image formation process in CMOS cameras is proposed
- Given multiple low-resolution (LR) images that are RS affected, a unified framework is developed to obtain an undistorted and super-resolved image by alternating between solving for the underlying high-resolution (HR) image and the row-wise motion
- Assumption: The first LR image is free from RS effect and has only undergone a downsampling operation with respect to the HR image

RS-SR Image Formation Model

- The classical SR equation for a CCD camera
  \[ g = D \cdot W f \]  
  \[ \epsilon \geq 1 \] : super-resolution factor
  \[ g \in \mathbb{R}^{M \times N} \] : LR image of size \( M \times N \) lexicographically ordered
  \[ f \in \mathbb{R}^{\epsilon M \times \epsilon N} \] : HR image of size \( \epsilon M \times \epsilon N \) lexicographically ordered
  \[ W \in \mathbb{R}^{\epsilon M \times \epsilon M} \] : warping matrix that multiplies \( f \) to produce its warped instance
  \[ D \in \mathbb{R}^{MN \times MN} \] : decimation matrix which averages \( \epsilon \) neighboring pixels in the HR image

- Image formation model for an RS camera - static versus moving

  ![Image formation model](image)

  1. Actual LR sensor plane
  2. Virtual HR sensor plane
  3. Scene plane

  The virtual HR sensor plane is the HR representation of the scene that an HR camera would have captured
  It is this HR image that is to be recovered
  For an SR factor of 2, a pair of rows in the HR plane experience the same motion
  For an SR factor of \( \epsilon \), this corresponds to a block of \( \epsilon \) rows in the virtual HR sensor plane having the same motion associated with them
  Unlike in a GS camera where all rows of \( W \) are associated with a single camera motion, in RS cameras, the motion varies depending on which particular block of rows in the HR image the pixel belongs to
  (1) can be rewritten for a CMOS camera as
  \[ g = D \cdot W f \]  
  where \( W \) is the warping matrix that multiplies \( f \) to produce an RS image
  There are \( M \) warps associated with \( W \) as against a single warp for \( W \)

Optimization Problem

- Aim: Recover \( f \) given \( K \) LR images \( \{g_k\} \), where \( g_k = D \cdot W \cdot f \), for \( k = 1 \) to \( K \)
- Alternating minimization scheme to solve for the two unknowns \( f \) and \( W \)
- The minimization sequence \( \{f_p, W_k\} \), where \( p \) indicates the iteration number, can be built by alternating between two minimization subproblems
- Starting with an initial estimate \( f_0 \) (obtained by upsampling the first LR image), the two alternating steps are: step 1) estimate \( W_k \) using the previous iterate \( f_{p-1} \), step 2) use the current estimate \( W_k \) to compute \( f_p \)

Warp Estimation

- Estimate a single camera pose/warp from a discrete camera pose space \( S \)
- The cost function is formulated such that a few camera poses around the centroid of these poses yields the true motion for that row
  \[ \hat{w}_{k_p} = \arg \min_{w_k} \{ ||g_k(i) - D \cdot W f||_2^2 + \lambda ||w_k(i)||_1 \} \]  
  subject to \( w_k(i) \geq 0 \)

- \( g_k(i) \) denotes the \( i \)-th row of the LR image \( g_k \) and \( w_k(i) \) is its corresponding weight vector of size \( |S| \times 1 \) which chooses the required set of poses from the search space \( S \)
- Since \( w_k(i) \) is sparse, \( l_1 \)-norm with non-negativity is imposed so as to choose a sparse set of camera poses with corresponding weights to calculate the centroid
- The weighted average of the rotations and translations in the search space is found to give the centroid pose; \( R_c = \hat{w}_{k_p} \circ \{ R_j \}_{j=1}^{|S|} \) and \( T_c = \hat{w}_{k_p} \circ \{ T_j \}_{j=1}^{|S|} \)
- The centroid of these poses yields the true motion for that row
- The weighted average of the rotations and translations in the search space is found to give the centroid pose
- The weighted average of the rotations and translations in the search space is found to give the centroid pose

HR Image Estimation

- \( f_p = \arg \min_{f} \sum_{k=1}^{K} || D \cdot W_k \cdot f - g_k ||_2^2 + \alpha f^T L f \)  
- \( L \) is the discrete form of the variational prior

Experiments

- Input LR images
- Output RS-free HR image

- LR patch scaled up by 2
- Bicubic interpolation
- Glasner et al.
- Yang et al.
- Zhu et al.
- Ringaby + TV-SAR
- Ringaby + CDSR
- Our HR patch
- LR patches scaled up by 2
- Our HR patch

http://www.ee.iitm.ac.in/~ee10d038/RSSR.html

ee10d038@ee.iitm.ac.in