Super-resolution and Optical Flow reliability fields

Oleg Maslov and Konstantin Rodyushkin*
Intel Corporation

Abstract

Accurate optical flow estimation of motion fields is crucial for video super-resolution algorithms. Existing algorithms for optical flow calculation may produce erroneous motion fields on complex dynamic scenes with multiple moving, occluding non-rigid objects. In this paper we propose to use so called optical flow reliability weights in order to reduce impact of erroneous motion vector estimates on quality of super-resolution. We propose method for estimation of reliability weights which is based on structural analysis of motion vector fields. We present results on real video sequences and demonstrate the advantages of the proposed methods compared to conventional optical flow based super-resolution methods.

Keywords: Image and video processing, Super-resolution, Optical Flow

1 Introduction

We consider problem of applying optical flow based super-resolution methods to real world video sequences. Such problems evolves in everyday demand for enhancing quality of low resolution video from hand-held mobile devices, low resolution obsolete video cameras or from video web services with high degree of compression. Calculating optical flow (OF) between frames often produces inadequate and erroneous estimates that in place results in bad quality of super-resolved images since registration information accuracy is crucial for super-resolution quality.

In this paper we propose a novel method for evaluating quality of optical flow data and then we apply this method in bayesian super-resolution framework and show the quality improvement on a set of real videos. The structure of this paper is organized as follows. In Section 2 we discuss papers where optical flow data is used as registration parameter for super-resolution. In section 3 we introduce our method. Then in Section 4 results for several video sequences are presented.

2 Related Work

There have been published a number of papers related to super-resolution based on OF registration data [Baker and Kanade 1999; Zhao and Sawhney 2002; Fransens et al. 2007]. Real video sequences may have multiple moving non-rigid and occluding objects. One method of describing non-rigid transformations between frames is a usage of dense motion vector fields proposed by Baker and Kanade in [Baker and Kanade 1999]. Authors proves feasibility of usage of OF as registration data for super-resolution. There exists a few number of methods for evaluating OF motion fields between images [Farnebäck 2003; Bruhn et al. 2005; Lucas and Kanade 1981]. However even most accurate of them may produce erroneous and noisy results. In [Lee and Kang 2003] authors considered registration error in global translation parameters. Using regularization term in Maximum a Posteriori (MAP) estimate they incorporate registration error caused by inaccurate motion information into minimization deconvolution functional. In [Zhao and Sawhney 2002] authors present OF error as additive stochastic noise in motion vectors measurements and show that consistency and accurateness of OF estimation is crucial to robust OF based super-resolution. In [Zomet et al. 2001] there was proposed robust median estimator in a super-resolution for dealing with outliers. It is used in gradient calculation during minimization of cost error function. It was shown that using median allows effectively suppress outliers caused by motion errors, noise, motion blur. In [Ben-Ezra et al. 2005] authors used adaptive super-resolution method, which detects blocks of 16×16 pixels with multiple motions present. In [Weiss 1998] it was proposed to Expectation-Maximisation (EM) algorithm for estimating simultaneously motion layers and motion parameters. In [Fransens et al. 2007] a set of so called visibility maps is introduced. Visibility maps signal whether or not a scene point on reconstructed high resolution (HR) image is visible in the low resolution (LR) input images. These maps are hidden binary variables, corresponding to visibility or occlusion, respectively. Expectation-Maximisation algorithm is then used, which iterates between estimating values for the hidden quantities, simultaneously optimizing the OF motion vectors and super-resolution image.

In this paper we propose to assess OF quality in form of set of so called OF reliability weights $W \in (0, 1)$. For each point of OF field between each pair of input LR images such OF reliability weights determine quality (accuracy) of OF in that point. These weights are estimated based on structural analysis of motion vector fields and are incorporated into super-resolution minimization functional. It allows us during deconvolution process to take into account reliable OF data and corresponding LR image pixels while rejecting OF errors like geometry and intensity outliers.

3 Method

This section discusses our method. We introduce imaging model and statistical bayesian super-resolution with reliability weights. Two subsections 3.2.1 and 3.2.2 describe methods we propose for estimating reliability weights.

3.1 Robust Super-Resolution with reliability weights

Suppose we have a set of $2N_f + 1$ low-resolution (LR) input images $T_i, i \in [-N_f, \ldots, N_f]$ with resolution $N_w \times N_h$ pixels. Imaging model similar to [Fransens et al. 2007] is used to reconstruct unknown high resolution (HR) image $J(\mathbf{x})$ of size $mN_w \times mN_h$ pixels, m - magnification factor so that

$$T_i(\mathbf{x}) = S * P * J(F_i(\mathbf{x})) + \varepsilon, \quad (1)$$

where HR image $J(\mathbf{x})$ is warped by OF based operator $F(\mathbf{x})$; $*$ - two-dimensional convolution operator; then point spread function P and downsample operator S are applied and measurement noise ε is added to produce set of LR input images T_i. ε is assumed to be normally distributed with zero mean and covariance Σ, $\mathbf{x}$ is vector of two-dimensional image coordinates, subsampling operator S is dependant on magnification factor m. Since imaging model (1) is linear we may combine warping, PSF and subsampling operator as follows

$$T_i(\mathbf{x}) = \mathbf{H}_i^T(\mathbf{x})\mathbf{J} + \varepsilon, \quad (2)$$

where $\mathbf{H}_i(\mathbf{x})$ is a column vector dependant on operators $F(\mathbf{x})$, P and S, $\mathbf{J}$ is a column vector of HR image intensities $J(\mathbf{x})$ rearranged in lexicographical order.

*e-mail:{oleg.maslov,konstantin.rodyushkin}@intel.com

Suppose we have function $W_i(\mathbf{x})$ determining OF vector quality for each pixel $\mathbf{x}$ of observed LR image T_i. If $W_i(\mathbf{x}) = 1$ then OF vector for $\mathbf{x}$ is good and can be used for super-resolution, if $W_i(\mathbf{x}) = 0$ then OF vector is bad and the pixel $\mathbf{x}$ should not be used in super-resolution. Let us call this function $W_i(\mathbf{x})$ function of *OF reliability weights*.

Using approach from [Nestares et al. 2006] and introduced function of OF reliability weights let us formulate deconvolution functional in form of bayesian MAP estimate. This functional consists of likelihood part and a priori part. For likelihood and a priori nonlinear M-estimators based robust function are used which models outliers for pixel values. In this paper we use Cauchy function $\rho(x)$ as a robust M-estimator.

$$\rho(x) = \frac{c^2}{2} \log\left(1 + (x/c)^2\right) \tag{3}$$

As an a priori probability for image we use Markov Random Fields (MRF) since they are a common choice to model prior probabilities on images. Resulting deconvolution functional is:

$$E(\mathbf{J}) = \sum_i \sum_{\mathbf{x}} W_i(\mathbf{x}) \rho\left(\frac{\mathbf{H}_i^T(\mathbf{x})\mathbf{J} - \hat{T}_i(\mathbf{x})}{\sigma_l}\right) + \lambda \sum_{(k,l)\in\Omega} \rho\left(\frac{\mathbf{V}_{(k,l)}^T \mathbf{J}}{\sigma_p}\right) \to \min \tag{4}$$

where $\hat{T}_i(\mathbf{x})$ are input LR images; σ_l, σ_p are scale factors defining error magnitude that is considered to be an outlier in likelihood and prior terms correspondingly; λ - relative weight given to the prior with respect to the likelihood; Ω - index pairs of neighbor pixels of MRF; $\mathbf{V}_{(k,l)}$ - is a vector defined by

$$\mathbf{V}_{(k,l)} = \{\delta_{i,k} - \delta_{i,l}\}, \quad i = 1, \ldots, mN_w \times mN_h \tag{5}$$

here $\delta_{i,l}$ - Kronecker symbol. For finding minimum of (4) we use nonlinear conjugate gradient optimization method [Barrett et al. 1994]. The gradient of (4) is

$$\nabla E(\mathbf{J}) = \sum_i \sum_{\mathbf{x}} \frac{W_i(\mathbf{x})}{\sigma_l} \rho'\left(\frac{\mathbf{H}_i^T(\mathbf{x})\mathbf{J} - \hat{T}_i(\mathbf{x})}{\sigma_l}\right) \mathbf{H}_i(\mathbf{x}) + \lambda \sum_{(k,l)\in\Omega} \frac{1}{\sigma_p} \rho'\left(\frac{\mathbf{V}_{(k,l)}^T \mathbf{J}}{\sigma_p}\right) \mathbf{V}_{(k,l)}. \tag{6}$$

This method uses iterations. Each iteration consists of the following steps:

1. Calculate gradient $\nabla E(\mathbf{J})$ using (6)
2. Calculate conjugate gradient search direction $\mathbf{D}$
3. Estimate search step size α along direction $\mathbf{D}$ by minimizing $E(\mathbf{J} + \alpha\mathbf{D})$ using Newton-Raphson method
4. Update HR image as $\mathbf{J} = \mathbf{J} + \alpha\mathbf{D}$

Iterations repeat until convergence.

3.2 Ways of calculating reliability weights

Super-resolution procedure described above implies that we have accurate OF motion fields $\mathbf{v}(\mathbf{x})$ which are used by warping operator $F(\mathbf{x}) = m(\mathbf{x} + \mathbf{v}(\mathbf{x}))$. Basically it means that motion vectors map pixels of arbitrary object of the scene exactly to the same pixels of this object across frames. But it might not be true because of the following reasons:

1. Some objects on the scene may disappear across frames. So there might be no correct OF vectors for pixels that belong to this object.
2. OF algorithm itself can produce erroneous results especially on boundaries of moving objects.
3. Objects of the scene can change appearance e.g. because of lighting conditions change or object deformation. In this case OF doesn't work well and even if we have good real OF vectors then these pixels are not good for super-resolution algorithm and should be rejected.

We propose two ways to estimate reliability weights for pixels based on OF map. Let us define reliability weight images for the two ways as W^1 and W^2 correspondingly. Here and below we will omit subscript image index because weights for each LR image are calculated independently and in the same way. The resulting reliability weights will be obtained as follows $W(\mathbf{x}) = W^1(\mathbf{x})W^2(\mathbf{x})$.

3.2.1 Normalization

The idea behind this way of reliability weight calculation is as follows. Each estimated pixel of HR image should use data from not more than one pixel from every input LR frame. e.g. let us consider OF map and corresponding warping operator $F(\mathbf{x})$ that warps two different pixels $\mathbf{x}_1$ and $\mathbf{x}_2$ into one pixel $\mathbf{x}_{hr}$. In other words

$$\mathbf{x}_{hr} = F(\mathbf{x}_1) = F(\mathbf{x}_2). \tag{7}$$

In this case super-resolution procedure considers intensities of both pixels $\mathbf{x}_1$ and $\mathbf{x}_2$ of the same input LR image as independent measurements which are used for estimation of one pixel $\mathbf{x}_{hr}$ of HR image $J(\mathbf{x})$. However it could not be true because from one LR frame we can get only one measurement for one HR pixel $\mathbf{x}_{hr}$. Then we should set such reliability weights w_1 and w_2 for LR pixels $\mathbf{x}_1$ and $\mathbf{x}_2$ so that $w_1 + w_2 = 1$ e.g. *normalize* them. We assume that x_1 and x_2 have equal impact on the HR pixel therefore we can set reliability weights for these pixels $w_1 = w_2 = 1/2$. In case of three pixels warped to one pixel of HR image we set reliability weights for such pixels as $1/3$ and so on. To obtain weights for the whole image we apply the following algorithm:

1. Calculate vector $\mathbf{C}_{hr}$ as follows:

   $$\mathbf{C}_{hr} = \sum_{\mathbf{x}'} \mathbf{H}(\mathbf{x}') \tag{8}$$

   where $\mathbf{C}_{hr}$ is vector of image C_{hr} pixels' values rearranged in lexicographical order and summation is done over all pixels of LR image. In result C_{hr} image is a "back projection" of LR image, which pixels are all set to one. Thus $C_{hr}(\mathbf{x})$ is the number of pixels from LR image used for estimation of the pixel $\mathbf{x}$ from HR image.

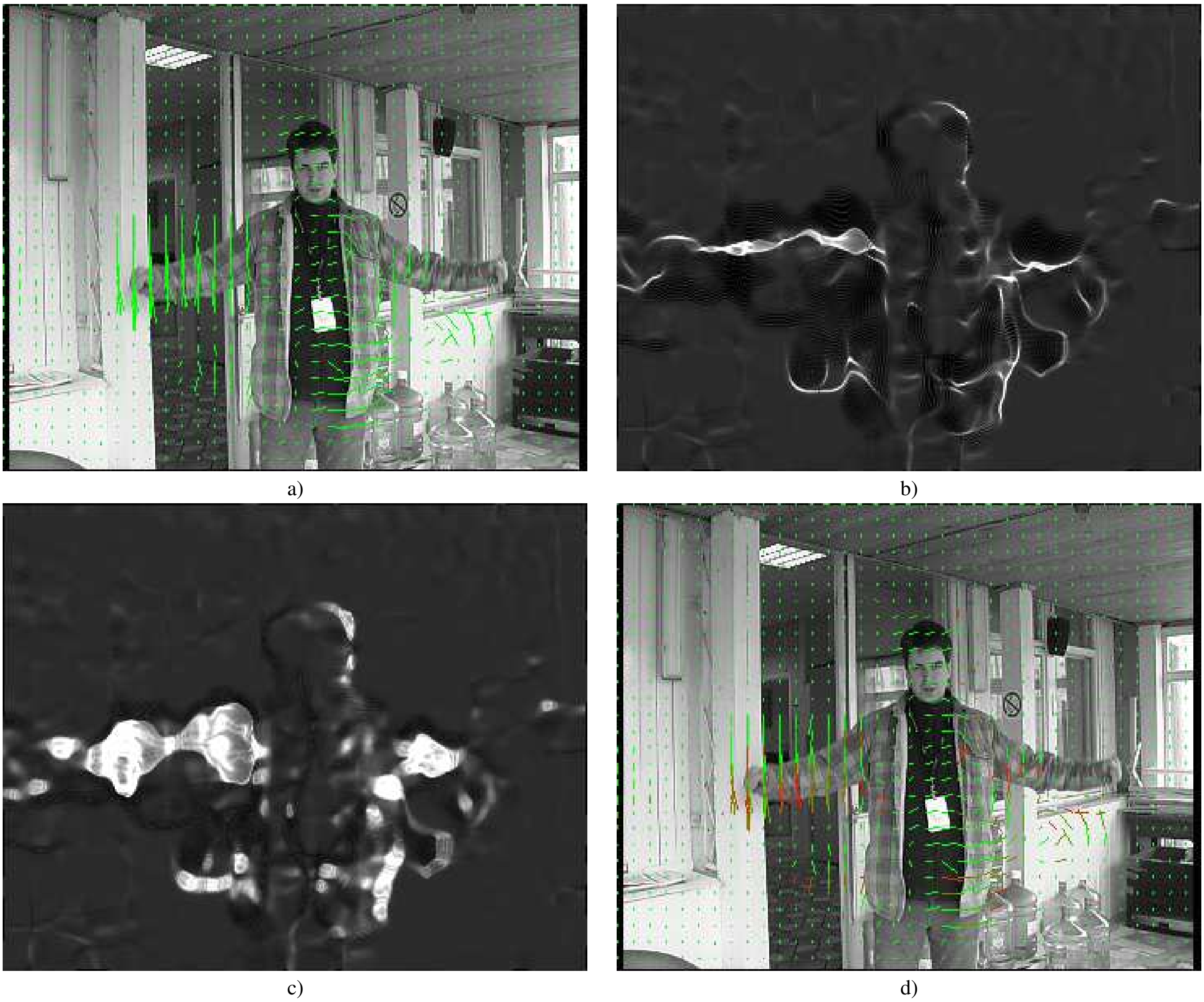
2. Map image C_{hr} to input LR image to obtain C_{lr} image ("direct projection")

   $$C_{lr}(\mathbf{x}) = \mathbf{H}^T(\mathbf{x})\mathbf{C}_{hr} = \mathbf{H}^T(\mathbf{x}) \sum_{\mathbf{x}'} \mathbf{H}(\mathbf{x}') \tag{9}$$

   In result $C_{lr}(\mathbf{x}) - 1$ is number of pixels from input LR image which are used to estimate HR pixel along with LR pixel $\mathbf{x}$.

3. Calculate weight image $W^1(\mathbf{x}) = \frac{1}{C_{lr}(\mathbf{x})}$

The illustration of such normalized transformation is shown on figure 1.



a) b) c) d)

Figure 1: *Calculation of OF reliability weights W^1. a) Input LR image with motion vectors, b) HR image C_{hr} - "inverse projection" of LR image containing one pixel values, c) C_{lr} image - "direct projection" of C_{hr}, d) Input LR image with reliability weighted motion vectors. Red vectors have zero weight and will not be used for super-resolution. Green vectors are related to good for super-resolution LR pixels with weight* 1.

3.2.2 Motion clusters

The second way of reliability weights calculation is based on the following idea. In order to obtain good super-resolution quality we should have image of the same object on several frames. It means that image of the object should have no significant deformations across frames. And therefore all motion vectors which belong to the object should be described by rigid motion e.g. translation, translation and rotation, or arbitrary affine transform.

This section describes algorithm of detecting groups of pixels with OF vectors within the same rigid motion i.e. *motion clusters*. When such detection is done we can mark all pixels which belong to the same rigid motion as good ones for super-resolution and mark the rest as bad and remove them from further consideration.

Suppose two images consist of pixels that either belong to one of N motion clusters described by rigid transformation $T(p_i, \mathbf{x})$, $i = 1 \ldots N$ or belong to motion cluster that could not be described by rigid transformation. Each rigid transformation is defined by parameter vector p_i. If pixel from first image belong to rigid transformed area i then its position on the second image can be calculated as follows $\mathbf{x} + T(p_i, \mathbf{x})$.

We introduce a set of hidden variables $\nu(\mathbf{x})$. If $\nu(\mathbf{x}) = 0$ then corresponding pixel x does not belong to any of rigid motion clusters. $\nu(\mathbf{x}) = 1, 2, \ldots, N$ means that pixel belongs to one of the N rigid motion clusters described by rigid transformation $T(p_i, \mathbf{x})$ and transformation parameters p_i ($i = 1, \ldots, N$).

We suppose that motion vectors are independent random values. So we can write joint probability distribution function (PDF) of all motion vectors $\mathbf{v}$ as product of PDF of each motion vector $\mathbf{v}(\mathbf{x})$ over all pixels $\mathbf{x}$.

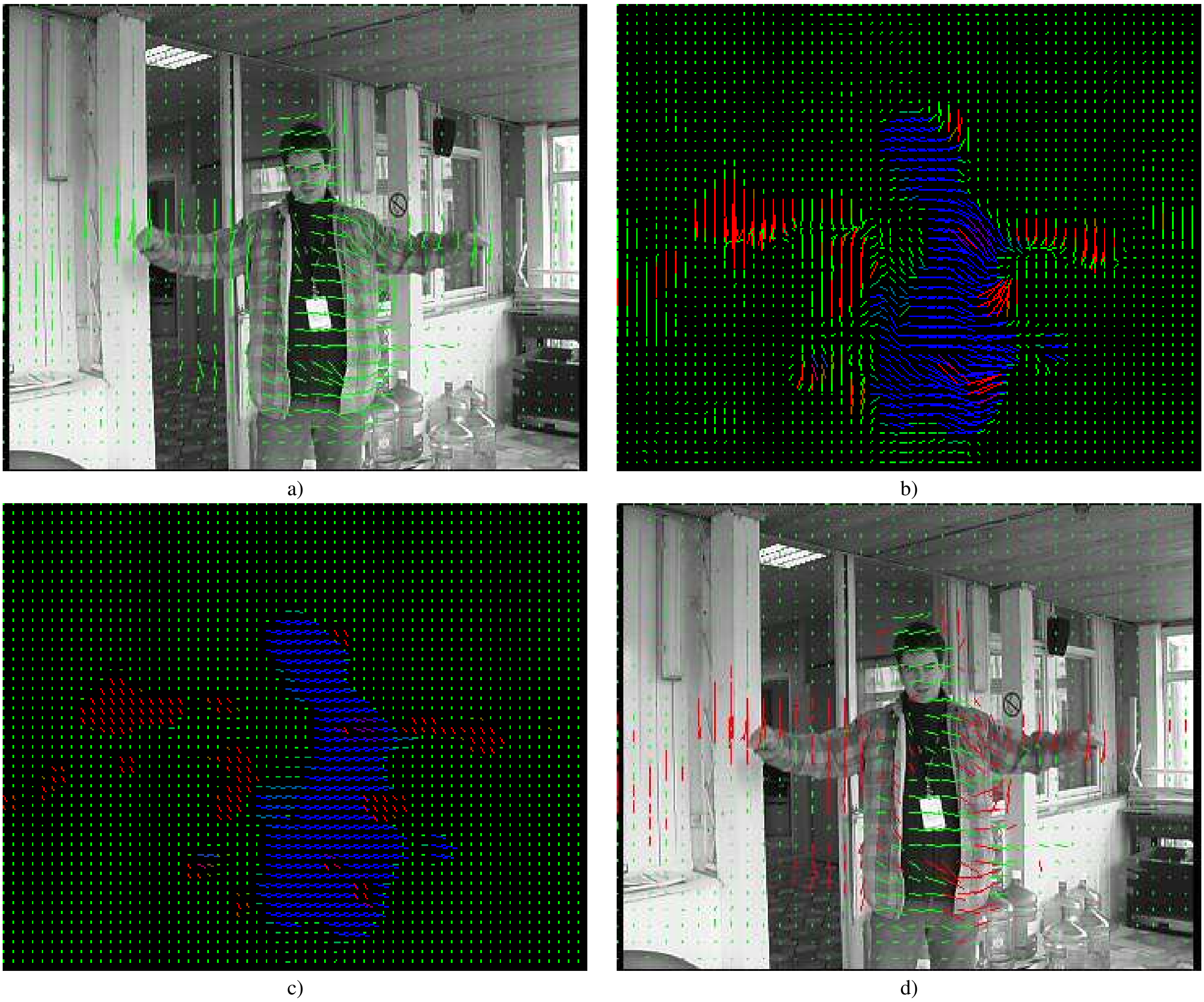


Figure 2: *Calculation of OF reliability weights W^2. a) Input LR image with motion vectors, b) Initial OF vectors. Red color denotes vector which does not belong to any rigid motion, green and blue colors indicate that vectors belong to first or second rigid motion cluster correspondingly. c) Rigid motion OF vectors. They show average motion directions for each cluster. d) Input LR image with reliability weighted motion vectors. Red vectors have zero weight and will not be used for super-resolution. Green vectors are related to good for super-resolution LR pixels with weight* 1.

$$P(\mathbf{v}|\mathbf{p},\nu) = \prod_{\mathbf{x}} P(\mathbf{v}(\mathbf{x})|\mathbf{p},\nu(\mathbf{x})), \tag{10}$$

here $\mathbf{p} = [p_0, p1, \ldots, p_N]$ is vector with transformation parameters , p_0 is a unknown parameters for non-rigid motion and $p_i, i = 1, \ldots, N$ are unknown parameters for motion vectors of i^{th} rigid transformation.

We suppose that the distribution function $P(\mathbf{v}(\mathbf{x})|\mathbf{p},\nu(\mathbf{x}))$ for each motion vector is described by 2D gaussian distribution:

1. If pixel $\mathbf{x}$ does not belong to rigid motion cluster i.e. $\nu(\mathbf{x}) = 0$ then

$$P\left(\mathbf{v}(\mathbf{x})|\mathbf{p},\nu(\mathbf{x}) = 0\right) = \frac{1}{2\pi\sigma_{out}^2} e^{-\frac{|\mathbf{V}(\mathbf{X}) - T(p_0,\mathbf{X})|^2}{0.5\sigma_{out}^2}} \tag{11}$$

2. If pixel $\mathbf{x}$ belongs to rigid motion cluster $k = 1 \ldots N$ i.e. $\nu(\mathbf{x}) = k$ then

$$P\left(\mathbf{v}(\mathbf{x})|\mathbf{p},\nu(\mathbf{x}) = k\right) = \frac{1}{2\pi\sigma^2} e^{-\frac{|\mathbf{V}(\mathbf{X}) - T(p_k,\mathbf{X})|^2}{0.5\sigma^2}} \tag{12}$$

where σ_{out} is deviation parameters for non rigid motion, and σ is deviation parameters for rigid motion. We suppose that σ_{out} and σ values are known and $\sigma_{out} \gg \sigma$.

The goal is to find transformation parameters of rigid and non rigid motions $\mathbf{p}$ using maximum likelihood criteria

$$\hat{\mathbf{p}} = \arg\max_{\mathbf{p}} \left\{ \log \sum_{\nu} P(\mathbf{v}|\mathbf{p},\nu) \right\} \tag{13}$$


k-3 k k+3

Figure 3: *Original 3 frames from set of 15 input frames.*

We will use expectation-maximization (EM) algorithm [Dempster et al. 1977] to solve problem (13).

E-step: One the E-step of the k^{th} iteration conditional probability distribution function $P(\nu|\mathbf{v}, \hat{\mathbf{p}}_{k-1})$ of hidden variables ν is estimated. $\hat{\mathbf{p}}_{k-1}$ is transformation parameters estimated on previous iteration during M-step. We suppose that $\nu(\mathbf{x})$ are independent from each other. Therefore we can rewrite the estimated distribution as product of PDFs of $\nu(\mathbf{x})$

$$\hat{P}(\nu|\mathbf{v}, \hat{\mathbf{p}}_{k-1}) = \prod_{\mathbf{x}} \hat{P}(\nu(\mathbf{x})|\mathbf{v}(\mathbf{x}), \hat{\mathbf{p}}_{k-1}), \tag{14}$$

$\hat{P}(\nu(\mathbf{x})|\mathbf{v}(\mathbf{x}), \hat{\mathbf{p}}_{k-1})$ is estimated as follows

$$\hat{P}\left(\nu(\mathbf{x}) = i|\mathbf{v}(\mathbf{x}), \hat{\mathbf{p}}_{k-1}\right) = w_i(\mathbf{x}), \tag{15}$$

$$w_i(\mathbf{x}) = \frac{P\left(\mathbf{v}(\mathbf{x})|\hat{\mathbf{p}}_{k-1}, \nu(\mathbf{x}) = i\right)}{\sum_{k=0}^{N} P\left(\mathbf{v}(\mathbf{x})|\hat{\mathbf{p}}_{k-1}, \nu(\mathbf{x}) = k\right)} \tag{16}$$

where $P(\mathbf{v}(\mathbf{x})|\hat{\mathbf{p}}_{k-1}, \nu(\mathbf{x}))$ are gaussian distributions defined in (11) and (12).

M-step: On the M-step transformation parameters $\mathbf{p}$ are estimated by means of optimization of following cost function

$$Q(\mathbf{p}) = \sum_{\nu} \hat{P}(\nu|\mathbf{v}, \hat{\mathbf{p}}_{k-1}) \log P(\mathbf{v}|\mathbf{p}, \nu) \tag{17}$$

Using (10), (11), (12), (14), and (16) it can be shown that optimization of (17) is equal to minimization of the following expressions

$$\hat{p}_{i,k} = \arg\min_{p_i} \left\{ \sum_{\mathbf{x}} w_i(\mathbf{x}) \left|\mathbf{v}(\mathbf{x}) - T(p_i, \mathbf{x})\right|^2 \right\} \tag{18}$$

here i is cluster index $i = 0 \ldots N$, $w_i(\mathbf{x})$ is estimated conditional PDF from (16).

For linear transformations T i.e. translation, scaling, rotation or arbitrary affine transform the task (18) is equal to solution of system of linear equations and can be solved analytically.

When new parameters $\hat{p}_{i,k}$ are calculated the next iteration is made. After several iterations we obtain the final estimation

$$\widetilde{p}_i = \hat{p}_{i,L} \tag{19}$$

where L is number of iterations.

Once we estimated parameters for N rigid motions clusters we can check if pixel belongs to one of the rigid motion clusters. In order to do this for each pixel $\mathbf{x}$ we calculate minimal distance between motion vector $\mathbf{v}(\mathbf{x})$ and motion vectors $T(\widetilde{p}_i, \mathbf{x})$ of rigid motion clusters found.

$$d(\mathbf{x}) = \min_{i=1\ldots N} |\mathbf{v}(\mathbf{x}) - T(\widetilde{p}_i, \mathbf{x})| \tag{20}$$

Based on this distance map we calculate weights for each pixel as follows

$$W^2(\mathbf{x}) = \begin{cases} 1, \text{if } d(\mathbf{x}) < \sigma \\ 0, \text{if } d(\mathbf{x}) \geq \sigma \end{cases} \tag{21}$$

The figure 2 shows example of rigid motion cluster estimation. In this example two rigid motion clusters and non-rigid motion cluster are detected.

4 Experiments

The purpose of experiment was to compare quality of conventional super-resolution algorithm and proposed super-resolution algorithm with optical flow reliability weights. The algorithm was

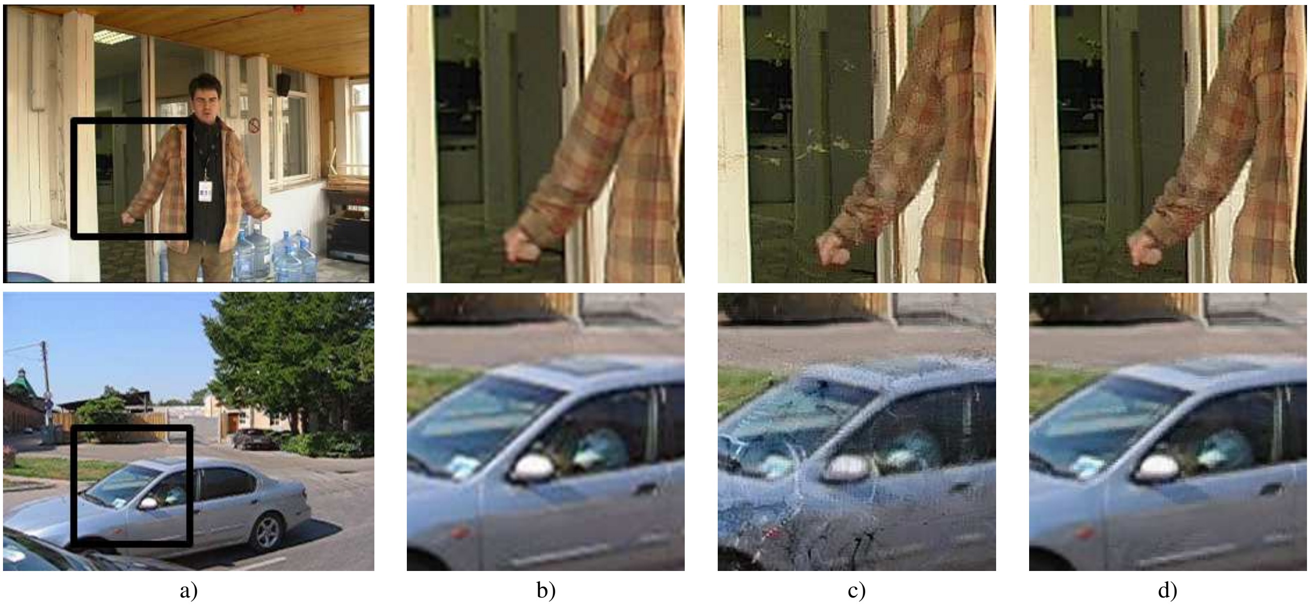


Figure 4: *Results of super-resolution algortihms: a) The original input LR image (central frame), b) result by bicubic interpolation, c) result by conventional super-resolution algorithm, d) result by super-resolution algorithm with reliability weights*

tested on two real world examples. There exists a variety of optical flow method. For estimating OF fields we used multiresolution two-frame OF algorithm proposed by Farnebäck in [Farnebäck 2003] since it gives a robust and adaptive to lighting conditions motion vectors estimation. OF averaging window size was 25×25, 3 levels of gaussian pyramids were used and 6 iterations were made on each pyramid level. Super-resolution was performed on the whole input image. For processing color images frames were transformed to luminance and color opponent space YCbCr. In this space image channels were super-resolved independently. Since luminance component contains high-frequency content typically it took at least $15 - 20$ iterations of conjugate gradient algorithm to converge. For color components we iterated 4 times. Magnification factor $m = 2$ was applied. For PSF function box filter of size $[m \times m]$ was used. We used 15 adjacent LR video frames to obtain each output HR video frame. Central frame magnified by factor m was used as an initial approximation for conjugate gradient algorithm. We assume here that typical scene contains 3 types of motion: foreground, background and non-rigid motion. So the number of motion clusters was set to 3. The variation parameters for OF vector PDFs (11), (12) were $\sigma = 2$ and $\sigma_{out} = 100\sigma$ correspondingly. Proposed algorithm was implemented in Microsoft Visual C/C++. Experiments were conducted on Pentium IV 3.2 Ghz machine.

For the first experiment we used real video captured by DV (digital video) camcorder in outdoor environment. Input image resolution was 360×288 pixels, 50 frames per second (FPS). Output video resolution Video contains static background and foreground with man rapidly waving arms. Some of the input 15 frames are presented in the top row of figure 3. The camera is shaking slightly in order to obtain sub-pixel shifts for super-resolution. The comparative results of bicubic interpolation and super-resolution algorithms are shown in the top row of figure 4. Here whole super-resolved image and zoomed fragment containing arms is shown. Rapidly moving arms produce large error in the optical flow estimation especially for the pairs of the most distant frames. It results in clearly visible ghost artifacts on the super-resolved image obtained by conventional algorithm. Our algorithm detects erroneous motion vectors and adjusts reliability weights correspondingly. In result output HR image contains significantly less artifacts caused by rapid motion.

For the second experiment video taken by Canon point and shoot digital camera. Input resolution was 320×240 pixels, FPS 15. The complex scene contains rapidly moving and occluding cars. The camera tracks car position. Some of input frames are presented in bottom row of figure 3. Results are shown in bottom row of figure 4. Conventional algorithm produces significant artifacts on moving cars while super-resolution with reliability weights suppresses them.

5 Summary

In this paper we proposed a novel method of using reliability weights as a robust way of producing high-quality super-resolution images based on erroneous and noisy estimates of motion vectors on complex dynamic videos. We proposed two novel methods for estimation of reliability weights based on structural analysis of motion vector fields. This approach allows to suppress visual artifacts' appearance during super-resolution caused by multiple moving deformable non-rigid objects, and may be applied also to scene detection where adjacent frames contain completely different scenes. Results on real video sequences demonstrate the advantages of the proposed methods compared to conventional optical flow based super-resolution method.

References

BAKER, S., AND KANADE, T. 1999. Super-resolution optical flow. *Technical Report CMU-RI-TR-99-36, The Robotics Institute, Carnegie Mellon University*.

BAKER, S., AND KANADE, T. 2002. Limits on super-resolution and how to break them. *IEEE Trans. Pattern Anal. Mach. Intell. 24*, 9, 1167–1183.

BARRETT, R., BERRY, M., CHAN, T. F., DEMMEL, J., DONATO, J., DONGARRA, J., EIJKHOUT, V., POZO, R., ROMINE, C., AND DER VORST, H. V. 1994. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. SIAM, Philadelphia, PA.

BEN-EZRA, M., ZOMET, A., AND NAYAR, S. K. 2005. Video super-resolution using controlled subpixel detector shifts. *IEEE Transactions on Pattern Analysis and Machine Intelligence 27*, 6, 977–987.

BRUHN, A., WEICKERT, J., AND SCHNORR, C. 2005. Lucas/kanade meets horn/schunck: Combining local and global optic flow methods. *IJCV 61*, 3 (February), 211–231.

DEMPSTER, A. P., LAIRD, N. M., AND RUBIN, D. B. 1977. Maximum likelihood from incomplete data via the em algorithm. *J. Royal Stat. Soc. 39*, 1–38.

FARNEBÄCK, G. 2003. Two-frame motion estimation based on polynomial expansion. *Proceedings of the 13th Scandinavian Conference on Image Analysis* (June-July), 363–370.

FARSIU, S., ROBINSON, D., ELAD, M., AND MILANFAR, P., 2004. Advances and challenges in super-resolution.

FRANSENS, R., STRECHA, C., AND VAN GOOL, L. 2007. Optical flow based super-resolution: A probabilistic approach. *CVIU 106*, 1 (April), 106–115.

LEE, E. S., AND KANG, M. G. 2003. Regularized adaptive high-resolution image reconstruction considering inaccurate subpixel registration. *IEEE Transactions on Image Processing 12*, 7, 826–837.

LUCAS, B., AND KANADE, T. 1981. An iterative image registration technique with an application to stereo vision. *IJCAI81*, 674–679.

NESTARES, O., HAUSSECKER, H., AND ETTINGER, S. 2006. Computing a higher resolution image from multiple lower resolution images using model-based, robust bayesian estimation. *U.S. Patent Application No. 20060002635*.

WEISS, Y. 1998. *Bayesian motion estimation and segmentation.* PhD thesis, MIT.

ZHAO, W., AND SAWHNEY, H. 2002. Is super-resolution with optical flow feasible? *ECCV02*, I: 599 ff.

ZOMET, A., RAV-ACHA, A., AND PELEG, S. 2001. Robust super-resolution. *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on 1*, I–645–I–650 vol.1.