Simultaneous Frame-rate Up-conversion of Image and Optical Flow Sequences

Shun Inagaki[1], Hayato Itoh[1] and Atsushi Imiya[2]
[1]*School of Advanced Integration Science, Chiba University, Yayoi-cho 1-33, Inage-ku, 263-8522, Chiba, Japan*
[2]*Institute of Management and Information Technologies, Chiba University, Yayoi-cho 1-33, Inage-ku, 263-8522, Chiba, Japan*

Keywords: Up-conversion, Optical Flow, Long Time Image-sequence, Variational Image Analysis.

Abstract: We develop a variational method for the frame-rate up-conversion of optical flow fields, in which we combine motion coherency in an image sequence and the smoothness of the temporal flow field. Since optical flow vectors define the motion of each point in an image, we can construct interframe images from low frame-rate image sequences using flow field vectors. The algorithm produces both interframe images and optical flow fields from a set of successive images in a sequence.

1 INTRODUCTION

In this paper, we develop a variational method for the frame-rate up-conversion of optical flow fields and video sequences. We combine image registration and optical flow computation for frame-rate up-conversion. Therefore, our algorithm produces both interframe images and optical flow fields from a set of successive images in a sequence.

For the application of well-established image analysis algorithms to low-frame-rate image sequences, which are common in bio-imaging and long-distance extrapolation, we are required to up-convert the frame rate of image sequences. In this conversion, the generation of interframe images is a fundamental requirement. Our proposed method generates the interframe images of a sequence to analyse motions in the sequence. The frame rate of images by long-distance extrapolation using a vehicle-mounted imaging system is low. To understand environment using optical flow field (Vardy and Moller, 2005) from such a low frame-rate image sequence, we are required to generate a temporally dense optical flow field, since well-established algorithms used to decipher images assume to use a high-frame-rate image sequence. To resolve this problem on the application of traditional method to use low-frame-rate image sequences, we are required to up-convert the optical flow field.

An optical flow field is a deformation field between two temporally successive frames. Therefore, the frame rate of an optical flow sequence depends on the frame rate of the input image sequence. The frame rate of a video sequence interpolated by interframe images is twice as high as that of the input video sequence. Therefore, increasing the frame rate of an optical-flow field sequence is required to up-convert the frame rate of the input image sequence by interpolating frames.

The methods of frame interpolation use optical-flow fields to generate inter-frame images. Several methods of frame-rate up-conversion for an image sequence using the optical flow field have been proposed (Werlberger et al., 2011; Lee et al., 2010). In this conversion, although for the generation of subframe images, the frame-rate up-conversion of an image sequence is a fundamental requirement, our objective is the up-conversion of the optical flow field sequence. However, our method generates subframe images of a sequence as intermediate information.

Superresolution recovers high-resolution images and/or image sequences from a low-resolution images and/or image sequences. There are some methods based on temporal coherence (Volz et al., 2011; Mobahi et al., 2012; Zimmer et al., 2011), which is based on the idea that successive frames have similar contents and small transformations (Sadek et al., 2009). Therefore, temporal coherence is effective for sequences of successive images. These method suppose a tube through frames in this assumption. In ref. (Volz et al., 2011; Zimmer et al., 2011), the authors proposed a method to calculate optical-flow fields by modelling the motion fields as being smooth along their trajectories. Furthermore, in ref. (Mobahi et al., 2012), the authors developed an interpolation method between two input images based on temporal coherence. Therefore, we introduce the assumption of tem-

Inagaki S., Itoh H. and Imiya A..
Simultaneous Frame-rate Up-conversion of Image and Optical Flow Sequences.
DOI: 10.5220/0005296800680075
In *Proceedings of the 10th International Conference on Computer Vision Theory and Applications* (VISAPP-2015), pages 68-75
ISBN: 978-989-758-089-5

poral coherence into optical flow temporal superresolution method. This implies we assume that the image sequences we calculate are temporally coherent.

Warping and morphing are fundamental techniques in computer graphics to interpolate and generate shapes and objects. In medical applications, morphing is used for the description of the deformation process of biological organs. This process predicts the deformable motion of biological organs in the human torso such as the beating heart, and the deformation of lungs the during breathing. Shape retrieval categorises and classifies shapes, and finds shapes from portions of shapes. In shape retrieval, the matching of shapes based on the diffeomorphism of shapes (Arrate et al., 2010) is used. In the matching process of images, the variational registration strategy (Modersitzki, 2004) is a typical tool.

In computer vision, a well-established evaluation method is the computation of least mean error between the ground truth and computed results. In up-conversion, there is, however, no ground truth for the evaluation of the up-converted results Therefore, we cannot evaluate results using tradition method in computer vision. For the evaluation of the up-converted optical flow field and images, we compute the temporal continuity of the field and image sequences. In this paper, we also prove the convergence condition for the simultaneous up-conversion of the optical-flow-field and image sequences. This condition conforms the uniqueness of the up-converted sequences.

2 INTER-FRAME OPTICAL-FLOW FIELD

Setting $f(\boldsymbol{x},t)$ to be a spatiotemporal image, we develop an algorithm to compute the optical-flow field $\boldsymbol{u}_{\frac{1}{2}}(\boldsymbol{x},t)$ of $f(\boldsymbol{x},t+\frac{1}{2})$. For the convenience of analysis, we set

$$f^+(\boldsymbol{x}) = f(\boldsymbol{x},t+1),\ f^-(\boldsymbol{x},t) = f(\boldsymbol{x},t), \quad (1)$$

$$g(\boldsymbol{x}) = f(\boldsymbol{x},t+\frac{1}{2}), \quad (2)$$

$$\boldsymbol{v} = \boldsymbol{u}_{\frac{1}{2}}(\boldsymbol{x},t),\ \boldsymbol{w} = \boldsymbol{u}_{\frac{1}{2}}(\boldsymbol{x},t+\frac{1}{2}). \quad (3)$$

Furthermore, stting

$$g(\boldsymbol{x}) = f^+(\boldsymbol{x}-\boldsymbol{w}),\ g(\boldsymbol{x}) = f^-(\boldsymbol{x}+\boldsymbol{v}),\ \boldsymbol{u} = \boldsymbol{v}+\boldsymbol{w}, \quad (4)$$

we can have the interframe image g and the up-converted optical flow fields, $\boldsymbol{v}$ and $\boldsymbol{w}$ as the minimisers of

$$J(g,\boldsymbol{v},\boldsymbol{w}) = I_+ + I_- + \alpha G + \beta U + \gamma V, \quad (5)$$

for [1]

$$I_+ = \int_{\boldsymbol{\Omega}} (g(\boldsymbol{x}) - f^+(\boldsymbol{x}-\boldsymbol{w}))^2 d\boldsymbol{x}, \quad (6)$$

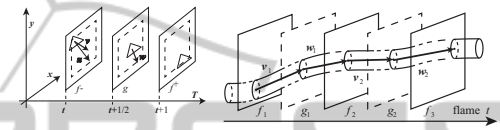
$$I_- = \int_{\boldsymbol{\Omega}} (g(\boldsymbol{x}) - f^-(\boldsymbol{x}+\boldsymbol{v}))^2 d\boldsymbol{x}, \quad (7)$$

$$G = \int_{\boldsymbol{\Omega}} |\nabla g|^2 d\boldsymbol{x}, \quad (8)$$

$$U = \int_{\boldsymbol{\Omega}} (|\nabla \boldsymbol{v}|^2 + |\nabla \boldsymbol{w}|^2) d\boldsymbol{x}, \quad (9)$$

$$V = |\boldsymbol{v}+\boldsymbol{w}-\boldsymbol{u}|^2, \quad (10)$$

where $|\boldsymbol{x}|$ is the l_2-norm of the vector $\boldsymbol{x}$. Figure 1 (a) shows the relationships of eq. (4). We extend eq. (5)



(a) (b)

Figure 1: Interframe optical flow. (a) g, f^-, f^+, $\boldsymbol{u}$, $\boldsymbol{v}$ and $\boldsymbol{w}$. (b) Input images f_1, f_2 and f_3, output images g_k and g_{k+1}, and output optical flow $\boldsymbol{v}_k$, $\boldsymbol{w}_k$, $\boldsymbol{v}_{k+1}$ and $\boldsymbol{w}_{k+1}$. The constant-intensity timeline tube is smooth along with the trajectories of optical-flow fields.

to an image sequence. Let the interframe image g_k be the image between two given frames

$$f_k := f(x,y,k),\ f_{k+1} := f(x,y,k+1). \quad (11)$$

Setting the vectors $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ to be the flow vector fields between f_k and g_k and between g_k and f_{k+1}, respectively, we assume that each $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ satisfy the optical flow constraint. Then, we have the energy functional

$$J(\{g_k,\boldsymbol{v}_k,\boldsymbol{w}_k\}_{k=1}^{n-1}) = \int_\Omega (J_d + \alpha J_g + \beta J_s + \gamma J_t) \mathrm{d}\boldsymbol{x}, \quad (12)$$

where

$$J_d = \sum_{k=1}^{n-1} ((g_k(\boldsymbol{x}) - f_k(\boldsymbol{x}+\boldsymbol{v}_k))^2 + (f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k) - g_k(\boldsymbol{x}))^2), \quad (13)$$

$$J_g = \sum_{k=1}^{n-1} |\nabla g_k|^2 d\boldsymbol{x} \quad (14)$$

$$J_s = \sum_{k=1}^{n-1} (|\nabla \boldsymbol{v}_k|^2 + |\nabla \boldsymbol{w}_k|^2), \quad (15)$$

$$J_t = J_t^{1st} + J_t^{2nd}, \quad (16)$$

[1] For vector-valued function $\boldsymbol{f}$ from $\mathbf{R}^2$ to $\mathbf{R}^2$, the vector gradient $\nabla \boldsymbol{f}$ is a 2×2 matrix. The norm $|\nabla \boldsymbol{f}|$ is the Frobenius norm of $\nabla \boldsymbol{f}$ which is defined as $\sqrt{tr \nabla \boldsymbol{f} \nabla \boldsymbol{f}^\top}$, where $tr\boldsymbol{F}$ is the trace of the matrix $\boldsymbol{F}$.

for

$$J_t^{1st} = \sum_{k=1}^{n-1}(|\boldsymbol{w}_k - \boldsymbol{v}_k|^2) + \sum_{k=1}^{n-2}(|\boldsymbol{v}_{k+1} - \boldsymbol{w}_k|^2) \quad (17)$$

$$J_t^{2nd} = \sum_{k=1}^{n-2}(|\boldsymbol{v}_{k+1} - 2\boldsymbol{w}_k + \boldsymbol{v}_k|^2 + |\boldsymbol{w}_{k+1} - 2\boldsymbol{v}_k + \boldsymbol{w}_k|^2). \quad (18)$$

Here, J_d is the data term for image registration. J_s, J_t and J_g are regularisation terms for the computation of smooth flow fields. The constraints J_t^{1st} and J_t^{2nd} require that the flow field is temporally piecewise smooth and temporally piecewise linear, respectively. As the minimisers of the variational problem, we obtain vector fields $\boldsymbol{v}_k$ and $\boldsymbol{w}_k$ and generate g_k simultaneously for the estimation of higher-frame-rate optical-flow fields and images.

3 NUMERICAL METHOD

3.1 Discretisation

To minimise the energy functional of eq. (12), setting

$$V_k = (g_k(\boldsymbol{x}) - f_k(\boldsymbol{x}+\boldsymbol{v}_k))\nabla f_k(\boldsymbol{x}+\boldsymbol{v}_k) + \gamma_1(2\boldsymbol{v}_k - \boldsymbol{w}_k - \boldsymbol{w}_{k-1}) + \gamma_2(\boldsymbol{v}_{k+1} - 4\boldsymbol{w}_k + 6\boldsymbol{v}_k - 4\boldsymbol{w}_{k-1} + \boldsymbol{v}_{k-1}), \quad (19)$$

$$W_k = (g_k(\boldsymbol{x}) - f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k))\nabla f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k) + \gamma_1(2\boldsymbol{w}_k - \boldsymbol{v}_k - \boldsymbol{v}_{k+1}) + \gamma_2(\boldsymbol{w}_{k+1} - 4\boldsymbol{v}_{k+1} + 6\boldsymbol{w}_k - 4\boldsymbol{v}_k + \boldsymbol{w}_{k-1}), \quad (20)$$

$$G_k = (2g_k(\boldsymbol{x}) - (f_k(\boldsymbol{x}+\boldsymbol{v}_k) + f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k))), \quad (21)$$

the Euler-Lagrange equations of eq. (12) are

$$\Delta\boldsymbol{v}_k - \frac{V_k}{\beta} = 0, \quad \Delta\boldsymbol{w}_k - \frac{W_k}{\beta} = 0, \quad \Delta g_k - \frac{G_k}{\alpha} = 0. \quad (22)$$

For the associate diffusion equations of the Euler-Lagrange equations of eq. (22),

$$\frac{\partial\boldsymbol{v}_k}{\partial t} = \Delta\boldsymbol{v}_k - \frac{1}{\beta}V_k, \quad \frac{\partial\boldsymbol{w}_k}{\partial t} = \Delta\boldsymbol{w}_k - \frac{1}{\beta}W_k, \quad \frac{\partial g_k}{\partial t} = \Delta g_k(\boldsymbol{x}) - \frac{1}{\alpha}G_k \quad (23)$$

using semi-implicit discretisation, we obtain the iteration form

$$(\boldsymbol{I} - \tau\boldsymbol{L})\boldsymbol{v}_k^{(m+1)} = \boldsymbol{v}_k^{(m)} - \frac{\tau}{\beta}V_k^{(m)}, \quad (\boldsymbol{I} - \tau\boldsymbol{L})\boldsymbol{w}_k^{(m+1)} = \boldsymbol{w}_k^{(m)} - \frac{\tau}{\beta}W_k^{(m)}, \quad (\boldsymbol{I} - \tau\boldsymbol{L})g_k^{(m+1)} = g_k^{(m)} - \frac{\tau}{\alpha}G_k^{(m)}, \quad (24)$$

where $\boldsymbol{L}$ and $\boldsymbol{I}$ are the discrete Laplacian matrix and the identity matrix whose degree depends on the size of the discrete images.

In each step of the iteration, the results are expressed on the Euler frame. Image samples obtained by the Lagrange frame do not guarantee correspondence between points. Therefore, we resample the results using the Lagrange frame[2]. In the Lagrange-frame-sampled images, we use Delaunay-triangle-based interpolation (Hjelle and Daehlen, 2006), since the method satisfies the minimum gradient property.

The iteration forms (Varga, 2000) derived in the previous sections are described in the form

$$(\boldsymbol{I} - \tau\boldsymbol{L})\boldsymbol{a}^{(m+1)} = \boldsymbol{a}^{(m)} - f(\boldsymbol{a}^{(m)}). \quad (25)$$

The Laplacian matrix $\boldsymbol{L}$ is described as

$$\boldsymbol{L} = \boldsymbol{D}\otimes\boldsymbol{I} + \boldsymbol{I}\otimes\boldsymbol{D}, \quad (26)$$

$$\boldsymbol{D} = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix},$$

where $\boldsymbol{A}\otimes\boldsymbol{B}$ is the Kronecker product of matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, assuming the Newmann condition on the boundary.

The eigenvalues of $\boldsymbol{D}$ are $\lambda_k = 4\sin^2\frac{\pi k}{2M}$ for the $M\times M$ matrix (Demmel, 1997), and the eigenmatrix (Strang and Nguyen, 1996) of $\boldsymbol{D}$ is

$$\boldsymbol{\Phi} = \left(\left(\varepsilon\cos\frac{(2j+1)i}{2\pi}\pi M\right)\right), \quad \varepsilon = \begin{cases} 1 & \text{if } j = 0 \\ \frac{1}{\sqrt{2}} & \text{otherwise.} \end{cases} \quad (27)$$

[2]There are two reference frames, which are called the Euler frame and Lagrange frame (Modersitzki, 2004). When we have an image B and an invertible transform φ, the frames are described as $B^{Lagrange}(\varphi(i,j)) := B(i,j)$ and $B^{Euler}(i,j) := B(\varphi^{-1}(i,j))$.

$\boldsymbol{\Phi}$ is the matrix of the DCT-II transform. Therefore, the matrix $\boldsymbol{L}$ is decomposed as

$$\boldsymbol{L} = (\boldsymbol{\Phi}\otimes\boldsymbol{\Phi})(\boldsymbol{\Lambda}\otimes\boldsymbol{I}+\boldsymbol{I}\otimes\boldsymbol{\Lambda})(\boldsymbol{\Phi}^\top\otimes\boldsymbol{\Phi}^\top) = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{U}^\top \tag{28}$$

and the eigenvalues of $\boldsymbol{L}$ are $\lambda_i+\lambda_j$. Then, since eq. (25) is redescribed as

$$\boldsymbol{a}^{(m+1)} = \boldsymbol{U}^\top(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{U}(\boldsymbol{a}^{(m)} - f(\boldsymbol{a}^{(m)})), \tag{29}$$

where

$$(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1} = Diag\left(\frac{1}{1-\tau(\lambda_i+\lambda_j)}\right), \tag{30}$$

that is, $\boldsymbol{a}^{(m+1)}$ is numerically computed from $\boldsymbol{a}^{(m)}$ using DCT-II (Strang and Nguyen, 1996) and a filtering operation for the Neumann boundary condition.

3.2 Convergence Conditions

3.2.1 Double Frame Method

For the case $n=2$, the interframe images g is computed from a pair of frames $f(\cdot,t)$ and $f(\cdot,t+1)$. From the third equation of (23), we have the iteration form

$$\boldsymbol{g}^{(m+1)} = \boldsymbol{U}(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{U}^\top(1-\frac{2\tau}{\alpha})\boldsymbol{g}^{(m)} + \boldsymbol{c}^{(m)}, \tag{31}$$

For a large m, we can assume that $\boldsymbol{c}^{(m)}$ is a constant vector. The iteration form

$$\boldsymbol{g}^{(m+1)} = \boldsymbol{U}(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{U}^\top(1-\frac{2\tau}{\alpha})\boldsymbol{g}^{(m)} + \boldsymbol{c}, \tag{32}$$

converges to a unique solution, if the relation

$$\max(|(1-\frac{2\tau}{\alpha})|) \le 1, \tag{33}$$

is satisfied. This condition derived from spectral radius of the matrix $(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}$ leads to the conclusion that the iteration form of eq. (32) produces a sequence $\boldsymbol{g}^{(m)}$ for $m\ge 1$ which converges a unique solution, if $\tau\le\alpha$.

For the up-conversion of the optical flow field, setting $\boldsymbol{c} = \frac{\gamma\tau}{\beta}(\boldsymbol{w}-\boldsymbol{u})$ and $\boldsymbol{\kappa} \ge \frac{\gamma\tau}{\beta}(\boldsymbol{f}^-(\boldsymbol{x}+\boldsymbol{v}) - \boldsymbol{g}(\boldsymbol{x}))\nabla_v\boldsymbol{f}^-(\boldsymbol{x}+\boldsymbol{v})$ we define the iteration form

$$\boldsymbol{v}^{(m+1)} = \boldsymbol{U}(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{U}^\top(1-\frac{\gamma\tau}{\beta})\boldsymbol{v}^{(m)} + \boldsymbol{c}+\boldsymbol{\kappa}. \tag{34}$$

for the copulation of $\boldsymbol{v}$.

Assuming $|\boldsymbol{f}^-(\boldsymbol{x}+\boldsymbol{v}) - \boldsymbol{g}(\boldsymbol{x})| \ll 1$ for a large m, and

$$\frac{\frac{\gamma\tau}{\beta}}{1-\frac{\gamma\tau}{\beta}} \ll 1, \tag{35}$$

for $\tau \ll \frac{\beta}{2\gamma}$ and $\tau \le \frac{2\beta}{\gamma}$ the iteration form converges. These conditions derives the conclusion that $\tau \ll \frac{\beta}{2\gamma}$ and $\tau \le \frac{2\beta}{\gamma}$ for convergence of the iteration form. The convergence of the computation of $\boldsymbol{w}$ is evaluated using the same iteration form with eq. (34).

3.2.2 Multiple Frame Method

Setting $\xi = -\frac{\tau}{\alpha}(f_k(\boldsymbol{x}+\boldsymbol{v}_k) + f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k))$ we have the iteration form

$$\boldsymbol{g}^{(m+1)} = \boldsymbol{U}(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{U}^\top(1-\frac{2\tau}{\alpha})\boldsymbol{g}^{(m)} + \xi, \tag{36}$$

Using the same methodology to evaluate the convergence condition for 5, we have the condition $\tau\le\alpha$.

Setting $\boldsymbol{X} = (\boldsymbol{v}_1,\boldsymbol{w}_1,\boldsymbol{v}_2,\ldots,\boldsymbol{v}_n,\boldsymbol{w}_n)$ we have the iteration form

$$\begin{aligned}\boldsymbol{X}^{(m+1)} &= \boldsymbol{U}(\boldsymbol{I}-\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{U}^\top\boldsymbol{X}^{(m)} \\ &\quad \boldsymbol{\Phi}^\top(\boldsymbol{I}-\frac{\tau}{\beta}((-\boldsymbol{\Lambda})+(-\boldsymbol{\Lambda})^2))\boldsymbol{\Phi}+\boldsymbol{\kappa},\end{aligned} \tag{37}$$

where

$$\boldsymbol{\kappa} = \begin{pmatrix} \frac{\tau}{\beta}(g_k(\boldsymbol{x}) - f_k(\boldsymbol{x}+\boldsymbol{v}_k))\nabla f_k(\boldsymbol{x}+\boldsymbol{v}_k) \\ \frac{\tau}{\beta}(g_k(\boldsymbol{x}) - f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k))\nabla f_{k+1}(\boldsymbol{x}-\boldsymbol{w}_k) \end{pmatrix}^\top. \tag{38}$$

Therefore, the spectral radius of matrix

$$\boldsymbol{M} = (\boldsymbol{I}-\frac{\tau}{\beta}((-\Lambda)+(-\Lambda)^2))\boldsymbol{\Phi}+\boldsymbol{\kappa}), \tag{39}$$

decides the convergence condition. The maximum value of the matrix $\boldsymbol{M}$ is established for $\lambda=\frac{1}{2}$. Furthermore, the relation

$$\max(|1-\frac{\tau}{\beta}((-\lambda)+(-\lambda)^2)|) = 1 \le 1. \tag{40}$$

is satisfied. These condition implies that the iteration form for flow computation yields a unique solution.

3.3 Multiresolution Method

For the up-conversion of the optical flow field, we are required to establish point correspondences from a large-displacement image sequence. Therefore, for the initial estimation of point correspondences we adopt the coarse-to-fine strategy using the pyramid transform. By establishing point correspondences in the higher level of the pyramid, we compute up-converted optical flow fields and the interframe image. as shown in Figure 2. Then, we estimate the point correspondences in the lower level in the pyramid by using the coarse-to-fine grid expansion.

Figure 2 illustrates a frame-rate up-conversion method with Gaussian pyramid transform. In this pipeline, By establishing point correspondences in the higher level of the pyramid, this pipeline first up-converts the optical flow fields and the interframe image. The recursively the algorithm estimates point correspondences on the lower level in the pyramid by using the coarse-to-fine grid expansion. In this algorithm, the number $frameNumber_{once}$ is selected as three for the method.

Algorithm 1: Coarse-to-fine inter-frame optical flow computation.

for $k = 1$ to $frameNumber_{all} - frameNumber_{once}$ **do**
  if $k = 1$ **then**
    for $j = 0$ to $frameNumber_{once} - 1$ **do**
      $\boldsymbol{v}_{k+j}, \boldsymbol{w}_{k+j}, g_{k+j} \leftarrow \emptyset$
    end for
  end if
  for $i = 0$ to *maxPyramidLevel* **do**
    Downsizing of f_ks to $2^{-i}\times$ Original Size
    Solve eq. (23)
    Upsizing of $\boldsymbol{v}_k$, $\boldsymbol{w}_k$ and g_k and setting as initial value
  end for
  output $\boldsymbol{v}_k, \boldsymbol{w}_k, g_k$
  for $j = 0$ to $frameNumber_{once} - 1$ **do**
    $\boldsymbol{v}_{k+j} \leftarrow \boldsymbol{v}_{k+j+1}$ as initial value.
    $\boldsymbol{w}_{k+j} \leftarrow \boldsymbol{w}_{k+j+1}$ as initial value.
    $g_{k+j} \leftarrow g_{k+j+1}$ as initial value.
  end for
end for

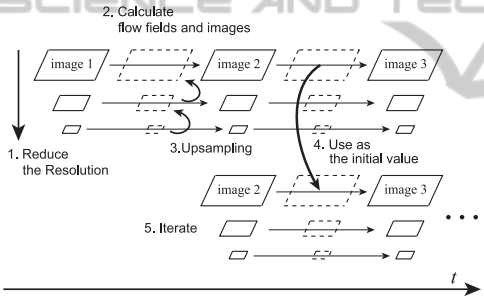
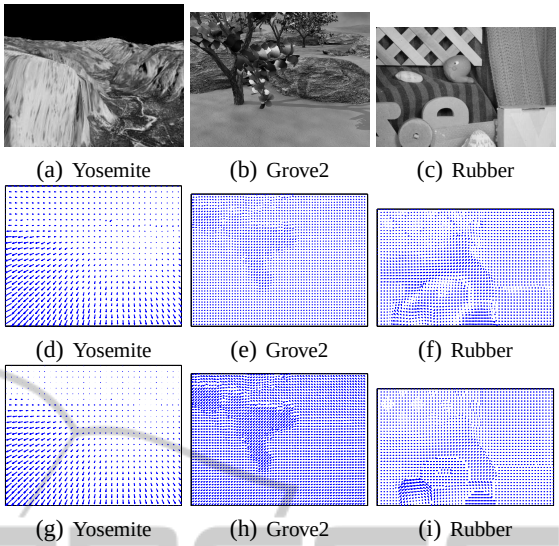


Figure 2: Multiframe image and optical flow up-conversion with the coarse-to-fine strategy. By establishing point correspondences in the higher level of the pyramid, the pipeline computes the up-covered optical flow fields and the inter-frame image. Then, the algorithm estimates point correspondences on the lower level in the pyramid by using the coarse-to-fine grid expansion.

4 NUMERICAL EXAMPLES

Figure 3 shows, from top to bottom, the images used, the flow fields obtained by our algorithm, and the ground-truth flow fields. In these examples, $f(x,y,t)$ and $f(x,y,t+2)$ are used as inputs. The sequences are, from left to right, Yosemite, Grove2 and RubberWhale. These results lead to the conclusion that the semicircular part and checkered part in RubberWhale are not calculated correctly. Since, in this example, there are no specific features, optical-flow field is diffused around.

Table 1 shows the errors between the computed flow fields and ground-truth flow fields according to



(a) Yosemite (b) Grove2 (c) Rubber
(d) Yosemite (e) Grove2 (f) Rubber
(g) Yosemite (h) Grove2 (i) Rubber

Figure 3: Results of the two-frame method. (a) An image from the Yosemite sequence. (b) An image from the Grove2 sequence. (c) An image from the RubberWhale sequence. (d) The subframe optical flow field of the Yosemite sequence. (e) The subframe optical flow field of the Grove2 sequence. (f) The subdrame optical flow field of the RubberWhale sequence. (g) The ground-truth optical flow field of the Yosemite sequence. (h) The ground-truth optical flow field of the Grove2 sequence. (i) The ground-truth optical flow field of the RubberWhale sequence.

the Middlebury evaluation page. Our method has two to three times the error of well-established methods. However, taking account of the fact that two flow fields and one image between two frames are calculated, our scores are in an acceptable range.

Figure 4 (a) shows five sucessive images from the Metronome sequence. Form this sequence, $f(x,y,t)$ and $f(x,y,t+2)$ are used as inputs for the evaluation of our algorithm. Furthermore, $f(x,y,t+1)$ is used as the ground truths. Smoothness and coherency of the subframe images and optical-flow fields are evaluated by the warp image error (WIE) and EMD such that

$$\mathrm{WIE}^2 = \int_{\mathbf{R}^2} |f(\boldsymbol{x}-\boldsymbol{w}_k, t+1) - f(\boldsymbol{x}+\boldsymbol{v}_k, t)|^2 d\boldsymbol{x}. \quad (41)$$

For the Metronom sequence, Fig. 4(b) shows, from top to bottom row, three frames from the input image sequence, the optical flow fields computed by the large-displacement optical flow computation, the subframe image sequence and its optical flow fields computed by the two-frame method and the subframe image sequence and its optical flow fields by the multiframe method using three frames.

Figure 5 shows histograms of the square norm of the optical flow vectors of three methods; the linear method (a), the two-frame method (b) and the three-frame method (c). For linear interpolation, we set

Table 1: Angular errors and endpoint errors in Grove2CRubberWhale and Yosemite sequences. These values are averages in one frame.

	EndpointError(pixel)		AngularError(degree)	
	mean	std	mean	std
Grove2($\boldsymbol{v}$)	1.08	1.03	16.62	17.76
Grove2($\boldsymbol{w}$)	1.05	1.02	16.03	17.76
RubberWhale($\boldsymbol{v}$)	0.61	0.73	20.05	22.92
RubberWhale($\boldsymbol{w}$)	0.59	0.73	19.48	22.92
Yosemite($\boldsymbol{v}$)	0.46	0.81	8.59	12.60
Yosemite($\boldsymbol{w}$)	0.47	0.86	9.17	14.32

$\boldsymbol{v} = \boldsymbol{w} = \boldsymbol{u}/2$. Figure 5(d) shows the average warp image errors. Figure 5 shows the three-frame method derives smooth high frame-rate optical flow fields by computing interframe images.

Figure 6 shows the original image for the beating heart sequence. Fig. 7 shows, from top to bottom row, shows three frames from the input image sequence, the optical flow fields computed by the large-displacement optical flow computation, the subframe image sequence and its optical flow fields computed by the two-frame method and the subframe image sequence and and its optical flow fields by the multi-frame method using three frames and the the subframe image sequence and its optical flow fields by 2 frame TV method [3]. Figure 8 shows comparison of the three-frame method and two-frame TV method. Top and bottom rows are results for the three-frame method and the two-frame TV method, respectively. From left to right interframe images, $\boldsymbol{v}$ and $\boldsymbol{w}$.

For the beating heart sequence, Fig. 9 shows histograms of the square norm of the optical flow vectors of four methods; the linear method (a), the two-frame method (b) and the three-frame method (c) and the two-frame method with TV constrain (d). Figures 9(e) is the average warp image errors for the beating heart sequence.

Figures 4 and 6 show that the interframe images are clearly generated. Furthermore, the smoothness and coherence of time trajectory of optical flow fields is evaluated by the earth movers' distance between histograms of norms of the optical flow vectors. Figures 5 and 9 show that the three-frame method produces smooth optical flow fields and smooth interframe images. Figure 8 implies that, for the beating heart sequence, the interframe optical flow fields computed using smoothness constrain produces are clear and accurate comparing the optical flow fields computed by using TV constrain.

Tables 2 and 3 list the parameters for numerical experiments. Furthermore, for the 2 frame TV method, we set $\varepsilon = 10^{-5}$.

Table 2: Parameters for the 2 frame method.

	α	β	γ
Card	8×10^{-2}	10^3	10^3
Metronome	8×10^{-2}	10^3	10^3

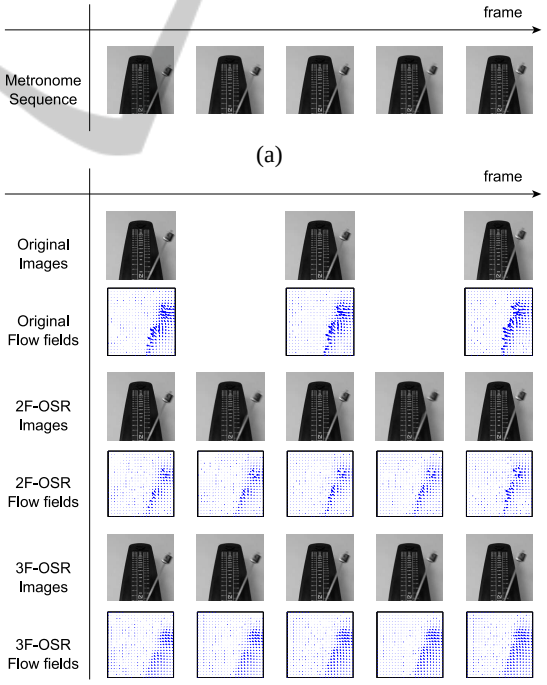
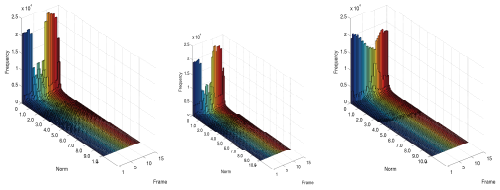
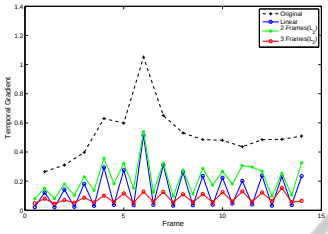


Figure 4: Up-conversion of the optical flow fields and interframe images. (a) is the original image sequence of the Metronome. (b) From the top to the bottom row, shows three frames from the input image sequence, the optical flow fields computed by the large-displacement optical flow computation, the subframe image sequence and its optical flow fields computed by the two-frame method and the subframe image sequence and its optical flow fields by the multi-frame method using three frames.

[3]The functional derivative of the total variation (TV) $\int_{\mathbf{R}^2} |\nabla f| d\boldsymbol{x}$ of f is $\nabla^\top \frac{\nabla f}{|\nabla f|}$ Numerically, $\nabla^\top \frac{\nabla f}{\sqrt{|\nabla f|^2+\varepsilon^2}}$, where ε is a small positive number, is achieved for the computation $\nabla^\top \frac{\nabla f}{|\nabla f|}$ By replacing $\int_{\mathbf{R}^2} |\nabla g|^2 d\boldsymbol{x}$ and $\int_{\mathbf{R}^2} |\nabla \boldsymbol{u}|^2 d\boldsymbol{x}$ in eq.(5) to $\int_{\mathbf{R}^2} |\nabla g| d\boldsymbol{x}$ and $\int_{\mathbf{R}^2} |\nabla \boldsymbol{u}|^2 d\boldsymbol{x}$, we define the two-frame TV method.

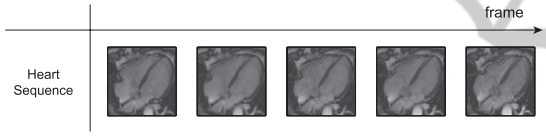


(a) Linear (b) 2frame (c) 3frame



(d) Warp image error of Metronome

Figure 5: Statistical comparisons of methods I. (a), (b) and (c) are histograms of least square norm of the optical flow field for the linear method, the two-frame method and the three-frame method, respectively, for the Metronome sequence. (d) is the average warp image errors for the Metronome, respectively.



(a)

Figure 6: Up-conversion of the optical flow fields and interframe images for the beating heart sequence. The original image sequence of the beating heart sequence.

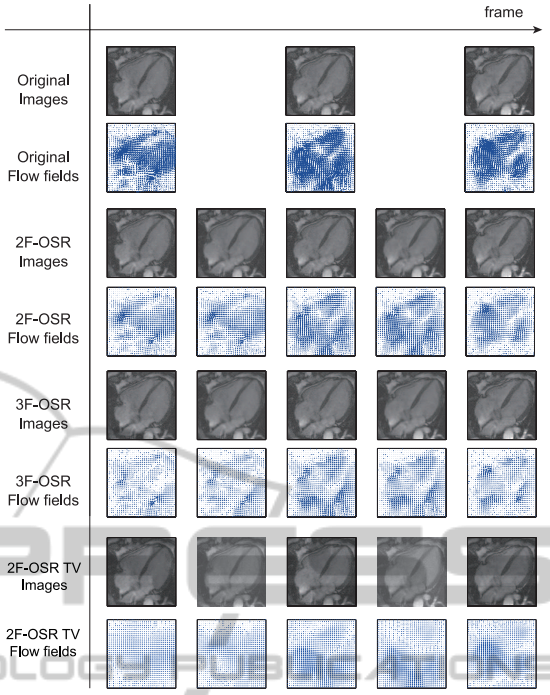
Table 3: 3 Parameters for 3 frame method.

	α	β	γ
Card	8×10^{-2}	3×10^{2}	3×10
Metronome	8×10^{-2}	10^{3}	10^{2}
Rubber Whale	7×10^{-2}	2×10^{2}	2×10
Grove 2	7×10^{-2}	2×10^{2}	2×10
Yosemite	7×10^{-2}	4×10^{2}	4×10

5 CONCLUSIONS

We developed a variational method for the frame-rate up-conversion of the optical-flow field-sequence in which we combine motion coherence in an image sequence and the smoothness of the temporal flow field.

For the evaluation of the accuracy of up-converted sequences, we proposed a results evaluation method without the ground truth. Furthermore, we have proved the convergence condition for the numerical method to up-convert the optical-flow-field and image



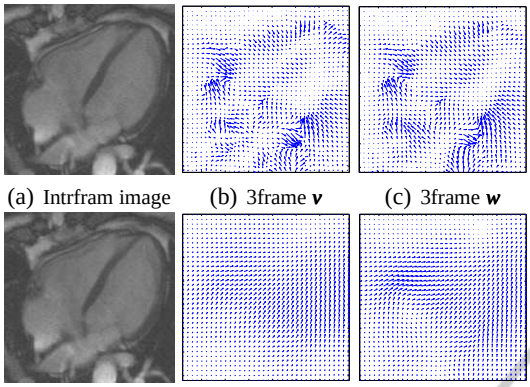
(a)

Figure 7: Up-conversion of the optical flow fields and interframe images for the beating heart sequence. From the top to the bottom row, shows three frames from the input image sequence, the optical flow fields computed by the large-displacement optical flow computation, the subframe image sequence and its optical flow fields computed by the two-frame method and the subframe image sequence and and its optical flow fields by the multi-frame method using three frames and the subframe image sequence and its optical flow fields by 2 frame TV method.

sequences to conform the uniqueness of the solutions. Since optical flow vectors define the motion of each point in an image, we can construct interframe images from low frame-rate image sequences using optical-flow field vectors. This implies that our method can be used to estimate ego-motion from a low frame-rate image sequence obtained by remote exploration.

ACKNOWLEDGEMENTS

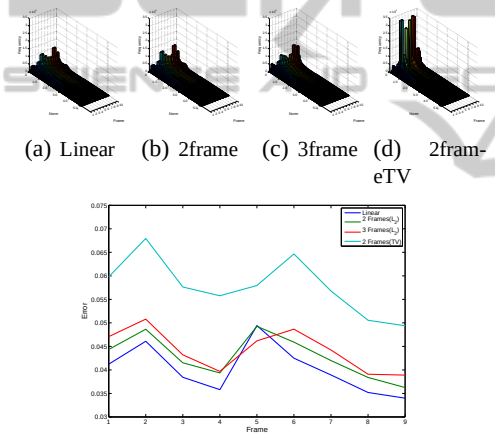
This research was supported by the “Computational Anatomy for Computer-Aided Diagnosis and Therapy: Frontiers of Medical Image Sciences” and “Multidisciplinary Computational Anatomy and Its Application to Highly Intelligent Diagnosis and Therapy” projects funded by a Grant-in-Aid for Scientific Research on Innovative Areas from MEXT, Japan, and by Grants-in-Aid for Scientific Research funded by the Japan Society for the Promotion of Science.



(a) Intrfram image (b) 3frame $\boldsymbol{v}$ (c) 3frame $\boldsymbol{w}$

(d) Interframe image (e) 2frame TV $\boldsymbol{v}$ (f) 2frame TV $\boldsymbol{w}$

Figure 8: Comparison of the three-frame and two-frame TV methods. Top and bottom rows are results for the three-frame method and the two-frame TV method, respectively. From left to right interframe images, $\boldsymbol{v}$ and $\boldsymbol{w}$.



(a) Linear (b) 2frame (c) 3frame (d) 2frameTV

(e) Warp image error

Figure 9: Statistical comparisons of methods II. (a), (b), (c) and (d) are histograms of least square norm of the optical flow field for the linear method, the two-frame method, the three-frame method, the two-frame method with TV constrain, respectively, for the Metronome sequence. (e) is the average warp image errors for the beating heart sequence.

REFERENCES

Arrate, F., Ratnanather, J. T., and Younes, L. (2010). Diffeomorphic active contours. In *SIAM J. Imaging Science*, volume 3, pages 176–198.

Demmel, J. W. (1997). *Applied Numerical Linear Algebra*. SIAM, Philadelphia.

Hjelle, O. and Daehlen, M. (2006). *Triangulations and Applications*. Springer, Heidelberg.

Lee, W. H., Choi, Y., Choi, K., and Ra, J. B. (2010). Frame rate up conversion via image fusion based on variational approach,. In *Prceedings of 17th IEEE ICIP*, pages 335–888. IEEE.

Mobahi, H., Collobert, R., and Weston, J. (2012). Deep learning from temporal coherence in video. In *Proceddings of VISAPP2012*, pages 367–372. SCITEPRESS.

Modersitzki, J. (2004). *Numerical Methods for Image Registration*. OUP, Oxford.

Sadek, R., Ballester, C., Garrido, L., Meinhardt, E., and Caselles, V. (2009). Deep learning from temporal coherence in video,. In *Proceddings of ICML09*, pages 733–744. ACM.

Strang, G. and Nguyen, T. (1996). *Wavelets and Filter Banks*. Wellesley-Cambridge Pres, Cambridge, MA.

Vardy, A. and Moller, R. (2005). Biologically plausible visual homing methods based on optical flow techniques. In *Connection Science*, volume 17, pages 47–89.

Varga, R. S. (2000). *Matrix Iterative Analysis*. Springer, Heidelberg, 2nd edition.

Volz, S., Bruhn, A., Valgaerts, L., and Zimmer, H. (2011). Modeling temporal coherence for optical flow,. In *13th ICCV*, pages 1116–1123. IEEE.

Werlberger, M., Pock, T., Unger, M., and Bischof, H. (2011). Optical flow guided tv-l1 video interpolation and restoration. In *LNCS*, volume 6819, pages 273–286.

Zimmer, H., Bruhn, A., and Weickert, J. (2011). Optic flow in harmony. In *IJCV*, volume 93, pages 368–388.