Video Completion by Motion Field Transfer

Takaaki Shiratori*   Yasuyuki Matsushita**   Sing Bing Kang† Xiaoou Tang**

* The University of Tokyo   ** Microsoft Research Asia   † Microsoft Research

東京大学
THE UNIVERSITY OF TOKYO

Microsoft Research Asia
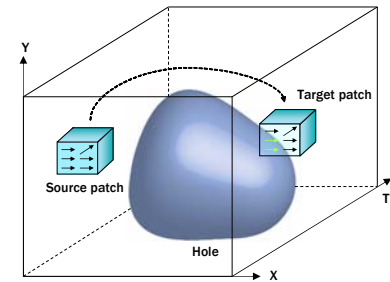
→ Why Use Motion Field?

− Color-based method : Requires similar <u>color</u> & <u>motion</u>
+ Motion-based method : Requires only similar <u>motion</u>

↓

More chance to fill-in a hole!

Motion can be copied from video portions with different appearance.

→ Algorithm Overview

1. **Pixel wise local motion estimation**
   - Lucas-Kanade Algorithm
2. **Motion field transfer by non-parametric sampling**
3. **Color propagation using the estimated motion field**

→ Motion Dissimilarity

$\mathbf{m} = (ut, vt, t)^T$: **motion vector in 3D spatio-temporal domain**
(**t** : frame interval)

Motion dissimilarity

$$d_m(\mathbf{m}_0, \mathbf{m}_1) = 1 - \frac{\mathbf{m}_0 \cdot \mathbf{m}_1}{|\mathbf{m}_0||\mathbf{m}_1|} = 1 - \cos\theta$$
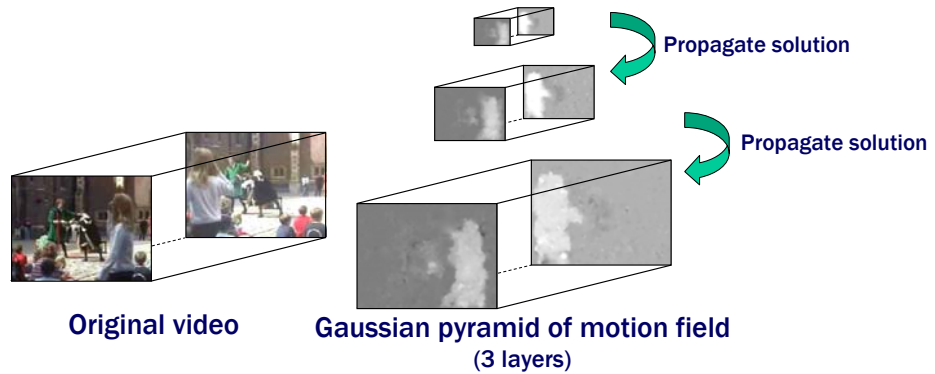
θ : **Angular distance between $\mathbf{m}_0$ and $\mathbf{m}_1$ in homogeneous coordinates**

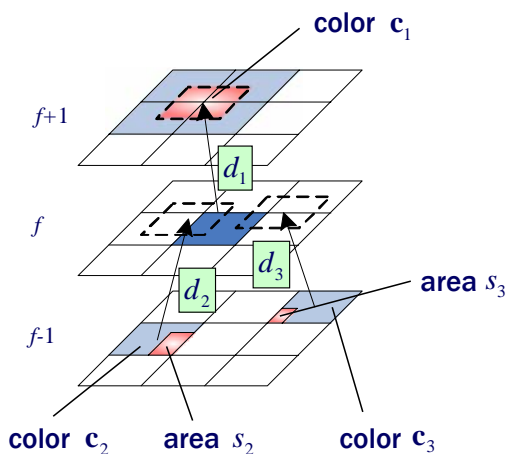## → Motion Field Transfer

An optimal source patch P_s minimizes

$$d(P_s(\mathbf{x}_s), P_t(\mathbf{x}_t)) = \frac{1}{|D|}\sum_{\mathbf{p}\in D} d_m(\mathbf{m}(\mathbf{x}_s + \mathbf{p}), \mathbf{m}(\mathbf{x}_t + \mathbf{p})) \; .$$

Find optimal source patches in coarse-to-fine manner.

Propagate solution

Propagate solution

Original video

Gaussian pyramid of motion field
(**3 layers**)

→ Color Propagation

Color of ▱ **is estimated from** ▱

color $\mathbf{c}_1$

$f+1$

f

d_1

d_3

d_2

area s_3

$f-1$

color $\mathbf{c}_2$ area s_2 color $\mathbf{c}_3$

Weighting factor for each ▱

$$w = \frac{s}{d}$$

$\left\{ \begin{array}{l} s \;: \text{ overlapped area} \\ d \;: \text{ patch distance} \end{array} \right.$

$$\Rightarrow \mathbf{c}(\text{▱}) = \frac{\sum w\mathbf{c}(\text{▱})}{\sum w}$$

For all hole pixels,

$$\mathbf{C} = \begin{bmatrix} \mathbf{W} \mid \mathbf{W}_b \end{bmatrix} \begin{bmatrix} \mathbf{C} \\ \mathbf{C}_b \end{bmatrix}$$

$\left\{ \begin{array}{l} \mathbf{C} \;: \text{ color of hole pixels} \\ \mathbf{C}_b : \text{ color of boundary pixels} \\ \quad\;\; \text{(already defined)} \end{array} \right.$

$$\mathbf{C} = \underbrace{(\mathbf{I} - \mathbf{W})^{-1}\mathbf{W}_b\mathbf{C}_b}$$

sparse & structurally symmetric matrix

→ Experimental Results

Hole-filling

Ground truth

Spatio-temporal hole (magenta) and computed optical flow (green)

Result of motion field transfer
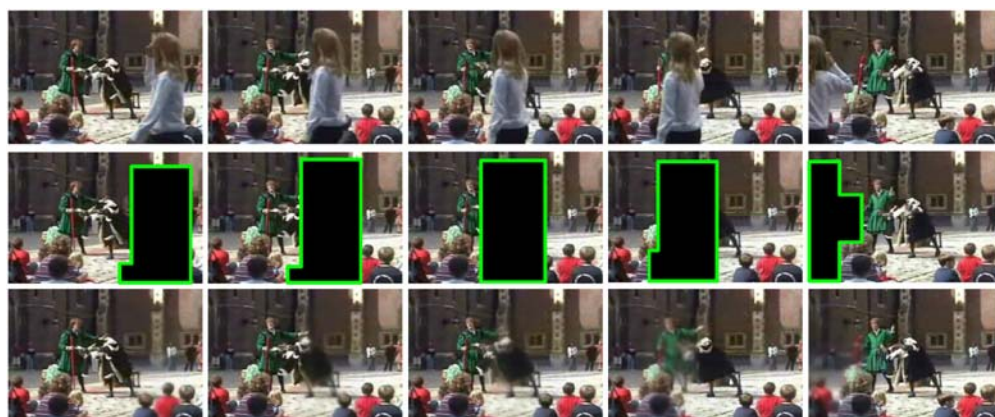
Result of video completion

Object removal

Original video

Foreground person is removed.

Completed video

Frame interpolation

⬜ : recovered frame

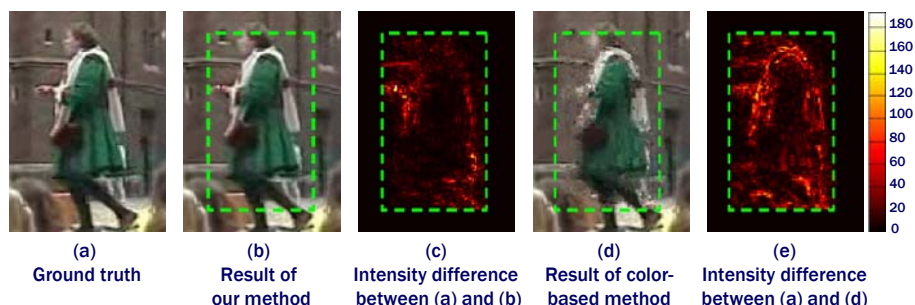### Frame rate recovery using the motion field prior

Missing frames are recovered by transferring the motion field of the prior video.
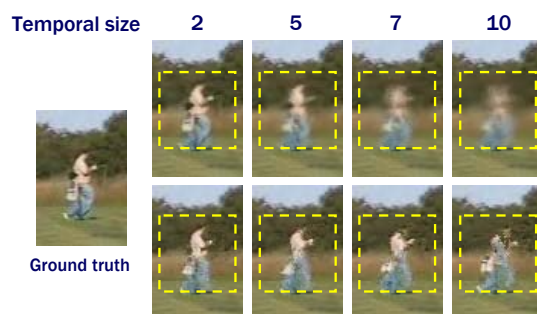
⬜ : recovered frame

Prior video

Output video

→ Discussion

Color-based method vs. Motion-based method

(a)
Ground truth

(b)
Result of our method

(c)
Intensity difference between (a) and (b)

(d)
Result of color-based method

(e)
Intensity difference between (a) and (d)

180
160
140
120
100
80
60
40
20
0

Blur effects

Temporal size   2   5   7   10

Ground truth

Upper : our method, lower : color-based method