Virtual View Rendering using Super-resolution with Multiview Images

Bumsub Ham
Digital Image Media Lab.
Yonsei University, Republic of Korea
Contents

• Introduction
• Virtual view rendering
• Motivation and overview
• Proposed method
  ▪ Stereo matching
  ▪ Geometric based forward warping
  ▪ Hole filling using backward warping
• Experimental results
• Conclusion
Introduction

• Feature of 3DTV system
  ▪ User interactivity
  ▪ 3D depth feeling

• Components of 3DTV system
  ▪ Ability of capturing 3D video
  ▪ Analysis & compression of multiview images
  ▪ Transmission of huge amount of data
  ▪ Display 3D video
Introduction

- An example of 3D TV system

Capture
Multi-view sequence

Multiview video encoder
Multiview depth video encoder
Multiview stereo matching

Transmission

Multiview video decoder
Multiview depth video decoder

Rendering / Coding
Freeview video rendering

Analysis / Coding

2D video (HDTV)
2D freeview video (Interactive HDTV)
3D freeview video (Renticular Parallax barrier)

Display
Virtual View Rendering

\[
x^v - x_c = f \frac{(x_i - x_c)B / d_i + T_x}{fB / d_i + T_z} = \frac{(x_i - x_c) + d_i \alpha_x}{1 + d_i \alpha_z / f}
\]

\[
y^v - y_c = f \frac{(y_i - y_c)B / d_i + T_y}{fB / d_i + T_z} = \frac{(y_i - y_c) + d_i \alpha_y}{1 + d_i \alpha_z / f}
\]

\[
(T_x / B, T_y / B, T_z / B) = (\alpha_x, \alpha_y, \alpha_z)
\]

\[
I_v(x, y) = \text{Vis}^L(x, y)(1 - \alpha)I_v^L(x, y) + \text{Vis}^R(x, y)\alpha I_v^R(x, y)
\]
Motivation and overview

• Motivation

Observation
When the virtual camera moves forward, the synthesized view is degraded.

In multiview configuration, we prevent the quality of a virtual view from being degraded using super-resolution concept.
Motivation and overview

- Overview
  - Stereo matching
  - Geometric based forward warping
  - Hole filling using backward warping
  - Deblur

Multiple images

Disparity Estimation

Input virtual viewpoint

Forward warping

Mapping to grid

Novel view synthesis

Deblur

Otherwise

Backward warping

Novel view synthesis
Proposed Method

- Stereo matching
  - Min’s algorithm

\[ e(p,d) = E(p,d) + n \]
\[ E^{k+1}(p) = \bar{e}(p) + \bar{E}^k(p) \]
\[ = e(p) + \lambda \sum_{m \in N(p)} w(p,m) E^k(m) \]
\[ = \frac{1 + \lambda \sum_{m \in N(p)} w(p,m)}{1 + \lambda \sum_{m \in N(p)} w(p,m)} \]

Proposed Method

- Geometric based forward warping
  - Forward warping: Multiple matching → hole problem
  - Backward warping

However, we do not use interpolation!

Forward is more suitable than backward warping
Proposed Method

- Geometric based forward warping

![Diagram of camera setup and virtual camera movement](image)
Proposed Method

- Geometric based forward warping

\[
I'(x', y') = I_i(x_i, y_i)
\]

\[
x' = x_c + \left( \frac{(x_i - x_c)B_i/d_i(x_i, y_i) + T_{i,x}}{fB_i/d_i(x_i, y_i) + T_{i,z}} \right)
\]

\[
y' = y_c + \left( \frac{(y_i - y_c)B_i/d_i(x_i, y_i) + T_{i,y}}{fB_i/d_i(x_i, y_i) + T_{i,z}} \right)
\]

Foreground object: large disparity
Proposed Method

- Geometric based forward warping
  - Disparity refinement
    - Disparity influences the quality of synthesized view.

\[
\begin{align*}
    x^v - x_c &= f \left( \frac{(x_i - x_c)B_y}{d_i(x_i, y_i)} + T_{i,x} \right) = \frac{(x_i - x_c) + d_i(x_i, y_i)\alpha_{i,x}}{1 + d_i(x_i, y_i)\alpha_{i,z}/f} \\
y^v - y_c &= f \left( \frac{(y_i - y_c)B_x}{d_i(x_i, y_i)} + T_{i,y} \right) = \frac{(y_i - y_c) + d_i(x_i, y_i)\alpha_{i,y}}{1 + d_i(x_i, y_i)\alpha_{i,z}/f}
\end{align*}
\]

\[
d_i(x_i, y_i) = \begin{cases} 
    d^0_i(x_i, y_i) + \sum_{k=i+1}^{G_i} d_k(x_k, y_k) \\
    x_k = x_{k-1} - d^0_{k-1}(x_{k-1}, y_{k-1}) \\
    d^0_i(x_i, y_i) + \sum_{k=G_i}^{i-1} d_k(x_k, y_k) \\
    x_k = x_{k+1} + d^0_{k+1}(x_{k+1}, y_{k+1})
\end{cases}
\]
Proposed Method

• Geometric based forward warping
  ▪ Two problem
    • Background disparity penetrates the foreground regions.
    • NOT one-to-one correspondence → Hole
Proposed Method

- Geometric based forward warping
  - Background disparity penetrates the foreground regions.
  - Depth ordering: pixels with largest disparity are visible.

However, this problem still occurs!
Proposed Method

- Geometric based forward warping
  - The reason why this problem occurs.
  - Similarity comparison: using neighboring disparity
Proposed Method

- Hole filling using backward warping
  - NOT one-to-one correspondence $\Rightarrow$ Hole
    - Backward warping with interpolated disparity map.
    - Geometric resampling: only used to change coordinate

\[
\begin{align*}
    x_i &= (x^v - x_c)(1 - f\alpha_{i,z} Dis_i(x^v, y^v)) + x_c + \alpha_{i,x} Dis_i(x^v, y^v) \\
    y_i &= (y^v - y_c)(1 - f\alpha_{i,z} Dis_i(x^v, y^v)) + y_c + \alpha_{i,y} Dis_i(x^v, y^v)
\end{align*}
\]

- Visibility function is needed.

\[
I^v(x^v, y^v) = I_i(x_i, y_i)V_i(x^v, y^v)
\]
Experimental Results : Exp1

- Conventional method
  - (from left to right) the virtual camera moves forward more.
Experimental Results : Exp1

- Proposed method
  - (from left to right) the virtual camera moves forward more.
Experimental Results: Exp1

• Conventional method
Experimental Results: Exp1

- Proposed method
Experimental results : EXP2

- A scene is synthesized by only applying backward warping.
- Not use interpolation

Backward warping based method

Proposed method

Virtual camera

Problem!!

Real camera

Virtual camera

Virtual camera pixel

Real camera pixel

Backward warped pixel
Experimental Results

- Complexity
  - Assume it takes \( N \) operation when one view is warped to the virtual camera coordinate
  - Conventional method: 11N
    - Disparity warping (2N)
    - Texture mapping (6N)
    - Blending (3N)
  - Proposed method: 12N (if we use 4 reference images)
  - Hole filling is ignored

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Tsukuba (ms)</th>
<th>Venus (ms)</th>
<th>Teddy (ms)</th>
<th>Cone (ms)</th>
<th>Robot (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional method</td>
<td>111.52</td>
<td>177.87</td>
<td>180.56</td>
<td>180.09</td>
<td>79.81</td>
</tr>
<tr>
<td>Proposed method</td>
<td>126.81</td>
<td>194.83</td>
<td>194.04</td>
<td>195.02</td>
<td>88.80</td>
</tr>
</tbody>
</table>
Experimental Results

- Objective evaluation: ‘Robot’ sequence
  - PSNR gain is mild (0.9dB) in comparison with visual quality.
    - Enhanced regions, that is, the foreground, occupy small portion.
    - Outliers - Disparity error.
Conclusion

• Summary
  ▪ Quality enhancement scheme
    • Prevent synthesized view from being degraded when the virtual camera moves forward
    • Apply SR concept to IBR
    • Forward and backward warping are used, properly

• Further work
  ▪ Investigate more elaborate data fusion
    • Robust to disparity error
  ▪ Overall PSNR gain is mild
    • Explore this in order to show good results both PSNR and visual quality
Thank You!
Any Questions?

http://diml.yonsei.ac.kr
mimo@yonsei.ac.kr
khsohn@yonsei.ac.kr