Corpus-based Visual Synthesis: An Approach to Artistic Stylization

Parag K. Mital\textsuperscript{1} \quad Mick Grierson\textsuperscript{1} \quad Tim Smith\textsuperscript{2}

\textsuperscript{1}Department of Computing, Goldsmiths, University of London
\textsuperscript{2}Department of Psychological Sciences, Birkbeck, University of London
Representation

- Perception must be supported by pre-attentive representations

- Abstract description of objects/scenes

- Numerous theories:
  - Gestalts
  - Geons (Biederman)
  - Object files (Wolfe)
  - Proto-objects (Rensink)
  - Indexicals (Pylyshyn)
  - Shapes (Marr)
  - Streams (Bregman)
Representation

- Perception must be supported by pre-attentive representations
- Abstract description of objects/scenes
- Numerous theories:
  - **Gestalts**
  - **Geons** (Biederman)
  - **Object files** (Wolfe)
  - **Proto-objects** (Rensink)
  - **Indexicals** (Pylyshyn)
  - **Shapes** (Marr)
  - **Streams** (Bregman)
Representation

- Perception must be supported by pre-attentive representations
- Abstract description of objects/scenes
- Numerous theories:
  - Gestalts
  - *Geons (Biederman)*
  - Object files (Wolfe)
  - Proto-objects (Rensink)
  - Indexicals (Pylyshyn)
  - Shapes (Marr)
  - Streams (Bregman)
**Representation**

- Perception must be supported by pre-attentive representations

- Abstract description of objects/scenes

- Numerous theories:
  - Gestalts
  - Geons (Biederman)
  - *Object files* (Wolfe)
  - Proto-objects (Rensink)
  - Indexicals (Pylyshyn)
  - Shapes (Marr)
  - Streams (Bregman)
Representation

- Perception must be supported by pre-attentive representations

- Abstract description of objects/scenes

- Numerous theories:
  - Gestalts
  - Geons (Biederman)
  - Object files (Wolfe)
  - Proto-objects (Rensink)
  - Indexicals (Pylyshyn)
  - Shapes (Marr)
  - Streams (Bregman)
Representation

Perception must be supported by pre-attentive representations

Abstract description of objects/scenes

Numerous theories:

- Gestalts
- Geons (Biederman)
- Object files (Wolfe)
- Proto-objects (Rensink)
- Indexicals (Pylyshyn)
- Shapes (Marr)
- Streams (Bregman)
Representation

* Artists are well aware of the role of abstract representations in perception

* Influence how we look at and where we look in a scene

* Art Movements
  * Impressionism
  * Pointilism
  * Cubism
  * Orphism
  * Expressionism
  * Abstract-Expressionism
Representation

- Artists are well aware of the role of abstract representations in perception

- Influence how we look at and where we look in a scene

- Art Movements
  - Impressionism
  - Pointilism
  - **Cubism**
  - Orphism
  - Expressionism
  - Abstract-Expressionism
Artists are well aware of the role of abstract representations in perception.

Influence how we look at and where we look in a scene.

Art Movements

- Impressionism
- Pointilism
- Cubism
- Orphism
- Expressionism
- Abstract-Expressionism
**Representation**

* Artists are well aware of the role of abstract representations in perception

* Influence how we look at and where we look in a scene

**Art Movements**

* Impressionism
* **Pointilism**
* Cubism
* Orphism
* Expressionism
* Abstract-Expressionism
Representation

- Artists are well aware of the role of abstract representations in perception
- Influence how we look at and where we look in a scene

Art Movements
- Impressionism
- Pointilism
- Cubism
- Orphism
- Expressionism
- Abstract-Expressionism
Representation

* Artists are well aware of the role of abstract representations in perception

* Influence how we look at and where we look in a scene

* Art Movements
  * Impressionism
  * Pointilism
  * Cubism
  * Orphism
  * **Expressionism**
  * Abstract-Expressionism
Representation

* Artists are well aware of the role of abstract representations in perception

* Influence how we look at and where we look in a scene

* Art Movements
  * Impressionism
  * Pointilism
  * Cubism
  * Orphism
  * Expressionism
  * Abstract-Expressionism
Related Work

- Filtering/Clustering
- Example-based Images
- Texture-transfer/Patch-based
- Dictionary methods/Collage approaches

Problem Statement

* Create automated artistic stylizations of images/videos using an understanding of the role of abstract representations in art and perception

* Allow for a range of styles through a simple set of parameters

* Needs to be fast in order to explore different styles quickly / run in real-time

Our approach

* **Build** corpus of abstract representations from user chosen images

* **Match** target image’s abstract representations to nearest ones in corpus

* **Synthesize** target image using closest matches

* **Interact** with a simple set of parameters effecting representation detection and synthesis
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
  - PhotoSynthesizer (iOS app)
- Conclusion
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
  - PhotoSynthesizer (iOS app)
- Conclusion
CBVS Framework

- **Build** corpus of abstract representations from user chosen images
- Match target image’s abstract representations to nearest ones in corpus
- Synthesize target image using closest matches
- Interact with a simple set of parameters effecting representation detection and synthesis

\[ \mathbf{R}_C = \{ R_1, R_2, ..., R_{NC} \} \]
Building the Corpus

* Need to represent:
  * Sparse/Dense strokes
  * Small/Large strokes

* Watershed?
* Posterization?
* Mean-Shift?
Building the Corpus

- Maximally Stable Color Regions
  - No need for multiple scale detections
  - Implicit ordering of regions
  - Simple set of parameters for discovering sparse/dense small/large strokes
  - Fast/Robust across multiple views (used in video tracking)
  - Similar process to the unconscious representations as theorized before
Increasing timesteps

= 

Denser detection

More expressive corpus
CBVS Framework

$R_C = \{ R_1, R_2, ..., R_{NC} \}$

$R_T = \{ R_1, R_2, ..., R_{NT} \}$

* Build corpus of abstract representations from user chosen images

* Match target image’s abstract representations to nearest ones in corpus

* Synthesize target image using closest matches

* Interact with a simple set of parameters effecting representation detection and synthesis
Matching

- Need to match regions in target to similar ones in corpus
- Describe Shape and Color
- Use Euclidean distance for shape values
- Use Perceptual distance for color values (CIEDE2000 formula)
- Nearest neighbor matching
Matching

\[ d_{R_i} = (\mu_{00}, \eta_{11}, \eta_{20}, \eta_{02}, L, a^*, b^* ) \]

\[ \eta_{ij} = \frac{\mu_{ij}}{\mu_{00} \left( 1 + \frac{i+j}{2} \right)} \]

- Need to match regions in target to similar ones in corpus
- Describe **Shape** and **Color**
- Use Euclidean distance for shape values
- Use Perceptual distance for color values (CIEDE2000 formula)
- Nearest neighbor matching
Matching

* Need to match regions in target to similar ones in corpus

* Describe Shape and Color

* Use **Euclidean distance** for shape values

* Use **Perceptual distance** for color values (CIEDE2000 formula)

* Nearest neighbor matching

\[ d(R_t, R_c) = d_s(R_t, R_c) + d_c(R_t, R_c) \]
CBVS Framework

- Build corpus of abstract representations from user chosen images
- Match target image’s abstract representations to nearest ones in corpus
- **Synthesize** target image using closest matches
- Interact with a simple set of parameters effecting representation detection and synthesis
Synthesis

From largest to smallest target region

\[ d(R_t, R_c) = d_s(R_t, R_c) + d_c(R_t, R_c) \]

- Find nearest neighbor
- Translate
- Rotate
- Scale
- Blend
Synthesis

\[ T = \text{centroid}_{R_t} - \text{centroid}_{R_c} \]

* From largest to smallest target region
  * Find nearest neighbor
  * Translate
  * Rotate
  * Scale
  * Blend
Synthesis

\[
\Theta = \frac{1}{2} \ast \arctan \left( \frac{2 \ast \frac{\mu_{11}}{\mu_{00}}}{\frac{\mu_{20}}{\mu_{00}} - \frac{\mu_{02}}{\mu_{00}}} \right)
\]

- From largest to smallest target region
- Find nearest neighbor
- Translate
- Rotate
- Scale
- Blend
Synthesis

\[ S_x = \frac{\text{width}_{RT}}{\text{width}_{RC}} \]

\[ S_y = \frac{\text{height}_{RT}}{\text{height}_{RC}} \]

- From largest to smallest target region
  - Find nearest neighbor
  - Translate
  - Rotate
  - Scale
  - Blend
Synthesis

- From largest to smallest target region
  - Find nearest neighbor
  - Translate
  - Rotate
  - Scale
  - Blend
Interaction

Discrete Parameters

- Spatial blending
- Temporal blending
- Motion tracking
- Timesteps
- Minimum region size
- Maximum region size
- Blending radius

Continuous Parameters
Spatial Blending
Temporal Blending
Temporal Blending
Timesteps

Increasing timesteps = Denser layers, More expressive
Minimum Size

Decreasing minimum size = Finer brush strokes
Blending Radius

Increasing radius = More of source texture
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
- Conclusion
Target Image

Corpus

Synthesis
Target Image  

Corpus  

Synthesis
Target Image

Corpus

Synthesis
Corpus-based Visual Synthesis: An Approach to Artistic Stylization

Target Image

Corpus
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
  - PhotoSynthesizer (iOS app)
- Conclusion
Memory Mosaicing

- Dynamic target (movie or webcam)
- Aggregate corpus over time using target, retaining only most recent N objects
- Only allow learning of objects with distance greater than threshold
Live demo
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
  - PhotoSynthesizer (iOS app)
- Conclusion
Augmented Reality Hallucinations

- Memory Mosaicing
- Uses Augmented Reality headset
- Exhibited during the Victoria & Albert Museum’s Digital Design Weekend co-located during the London Design Festival, 15,000 participants
- Short questionnaire for participants rating their experience (21 participants only)
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
  - PhotoSynthesizer (iOS app)
- Conclusion
PhotoSynthesizer

* Free iOS app allows user to synthesize target image
* No interaction besides selecting target and corpus
* Reveals synthesis process as painting regions over time
* Reached Top 50 app in Photo & Video in many countries
### PhotoSynthesizer

<table>
<thead>
<tr>
<th>Country</th>
<th># of countries - rank 10 reached</th>
<th># of countries - rank 100 reached</th>
<th># of countries - rank 500 reached</th>
<th># of countries - rank 1000 reached</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qatar</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Pakistan</td>
<td>-</td>
<td>-</td>
<td>26</td>
<td>26</td>
</tr>
<tr>
<td>Bermuda</td>
<td>-</td>
<td>-</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>Argentina</td>
<td>-</td>
<td>-</td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td>Philippines</td>
<td>-</td>
<td>-</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td>Belgium</td>
<td>-</td>
<td>-</td>
<td>43</td>
<td>43</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>-</td>
<td>-</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>United States</td>
<td>-</td>
<td>-</td>
<td>46</td>
<td>46</td>
</tr>
<tr>
<td>Spain</td>
<td>-</td>
<td>-</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td>Ireland</td>
<td>-</td>
<td>-</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td>India</td>
<td>-</td>
<td>-</td>
<td>123</td>
<td>123</td>
</tr>
<tr>
<td>France</td>
<td>-</td>
<td>-</td>
<td>128</td>
<td>128</td>
</tr>
<tr>
<td>Hungary</td>
<td>-</td>
<td>-</td>
<td>143</td>
<td>143</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Overall</th>
<th>Photo and Video</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td></td>
</tr>
<tr>
<td>Qatar</td>
<td>397</td>
</tr>
<tr>
<td>Pakistan</td>
<td>Nov 3, 2012</td>
</tr>
<tr>
<td>Bermuda</td>
<td>10</td>
</tr>
<tr>
<td>Argentina</td>
<td>25</td>
</tr>
<tr>
<td>Philippines</td>
<td>39</td>
</tr>
<tr>
<td>Belgium</td>
<td>41</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>39</td>
</tr>
<tr>
<td>United States</td>
<td>39</td>
</tr>
<tr>
<td>Spain</td>
<td>43</td>
</tr>
<tr>
<td>Ireland</td>
<td>64</td>
</tr>
<tr>
<td>India</td>
<td>123</td>
</tr>
<tr>
<td>France</td>
<td>128</td>
</tr>
<tr>
<td>Hungary</td>
<td>143</td>
</tr>
</tbody>
</table>
Outline

- CBVS Framework
- Results for Images/Videos
- Extensions
  - Memory Mosaicing
  - Augmented Reality Hallucinations
  - PhotoSynthesizer (iOS app)
- Conclusion
Conclusion

* Simple shape representation affords range of stylizations and a range of non/real-time applications

* Expressive control in a few parameters

* Future?
  * Better method of evaluation
  * Better metrics for shape description
  * Better temporal coherence