EUROGRAPHICS 2015 / O. Sorkine-Hornung and M. Wimmer
(Guest Editors)

Volume 34 (2015), Number 2

Guiding Image Manipulations using Shape-appearance Subspaces from Co-alignment of Image Collections

Chuong H. Nguyen[1,2] Oliver Nalbach[1] Tobias Ritschel[1,2] Hans-Peter Seidel[1,2]

[1]MPI Informatik [2]Saarland University, MMCI

1. Introduction

In this document, we provide further details on our submission "Guiding Image Manipulation Using Automatic Shape-Appearance Subspaces", such as

- Notation tables (Sec. 2),
- User study details (Sec. 3),
- Additional alignments for some classes shown in the paper (Sec. 5),
- Examples for shape and appearance transfer matrices (Sec. 6),
- Appearance transfer results using different methods and an analysis of the locally-rigid appearance transfer with different patch size (Sec. 7),
- Several complex images manipulations using our subspace (Sec. 4).

2. Table of Notations

Tables 1 and 2 show the sets, constants and general symbols used in alignment and subspace manipulation respectively.

Symbol	Meaning
	General:
a,b	Image a and b
	Alignment graph:
$d(a,b)$	Distance metric between images a and b
s	Neighborhood size of the min-pooling
ε	A small constant value to penalize longer paths
A	Pairwise distance matrix $\mathsf{A}_{a,b} = d(a,b)$
	Alignment:
$\mathbf{r}$	Searching neighborhood size
$f_{b,a}(\mathbf{x})$	The best corresponding pixel in a for pixel $\mathbf{x}$ in b
$E^{\text{dat}}_{b,a}(\mathbf{x})$	The data term of the alignment cost
l	Patch size of the data term
w_s	A spatial smoothness weighting function
σ_s	The spatial smoothness parameter of the data term
$E^{\text{mag}}_{b,a}(\mathbf{x})$	The flow magnitude term of the alignment cost
w_{mag}	Weight for the flow magnitude term
$E_{b,a}(\mathbf{x},\mathbf{y})$	The energy function of the pair $\mathbf{x}$ in b and $\mathbf{y}$ in a
$\Delta_{b,a}(\mathbf{x})$	Flow vector at $\mathbf{x}, \Delta_{b,a}(\mathbf{x}) =: \Delta(\mathbf{x})$
$c_{b,a}(\mathbf{x})$	The confidence of the flow $\mathbf{x}$ in b, $c_{b,a}(\mathbf{x}) =: c(\mathbf{x})$
κ	The curvature of $E_{b,a}(\mathbf{x}, f_{b,a}(\mathbf{x}))$
γ	A parameter to control the agreement check
$g(\mathbf{x})$	The blurred correspondence field
$w(\mathbf{x},\mathbf{y})$	The modified blur kernel
σ_d	The spatial smoothness parameter of w
σ_c	The steepness of the confidence function in w
$h(\mathbf{x})$	The locally-rigid regularization of $g(\mathbf{x})$

Table 1: *Table of notations for alignment*

3. User Study

3.1. Experimental Setup

The study was conducted on a desktop setup using a 24" LCD graphics display for visual stimuli. The participants were seated at approximately 60 cm distance to the screen and we asked them to place the mouse on the table such that they felt comfortable.

3.2. Manipulation Study

3.2.1. Participants

One female and six male subjects participated in the study. Our subjects were between 23 and 30 years old. All of them were graduate students in engineering or art. None of them reported to have issues with color perception. First, the subjects were introduced to the HSV color picker and interacting with it. Before performing our tasks, they were also given a training session to learn how to manipulate appearance and shape using both ours and the reference methods, until they felt comfortable to begin the study. The images used for the

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd. Published by John Wiley & Sons Ltd.

Symbol	Meaning
	Subspace construction:
n	Number of pixels
m	Number of basis vectors
α	A weight to control the contribution of shape/app.
$\{\mathbf{b}_j\}_{j=1}^{m}$	The set of basis vectors
λ_j	The eigen value corresponding to the basis vector $\mathbf{b}_j$
	Shape and appearance manipulation:
u	An image for manipulation
i	An iterator over the number of pixels n
j	An iterator over the number of basis vectors m
$\mathbf{v}$	A manipulated image vector
$\bar{\mathbf{v}}$	The closest projection of $\mathbf{v}$ in the subspace
$\mathbf{x}$	The coordinate of $\bar{\mathbf{v}}$ in the subspace
$\mathbf{k}$	The spatial weighting vector for subspace reconstruction
μ	A scalar regulating the prior's contribution
	Appearance transfer:
i	An iterator over the number of pixel in the query image u
j	An iterator over a square patch around i
$\mathbf{p}_j$	The source color at pixel j
$\mathbf{q}_j$	The target color at pixel j
$\mathbf{w}_j$	Weight depending on the distance between j and i and the alpha channels
$\bar{\mathbf{p}}$	The weighted centroid of the $\mathbf{p}_j$
$\bar{\mathbf{q}}$	The weighted centroid of the $\mathbf{q}_j$
R_i	The rotational component of the color transformation at i
$\mathbf{t}_i$	The translation component of the color transformation at i
δ	A regularization parameter
	Shape and appearance suggestions:
n_d	The number of directions used for suggestion

Table 2: *Table of notations for subspace construction and manipulation*

manipulation tasks did not include images used in the training session. We gave the subjects a short break after each task.

3.2.2. Appearance Manipulation Task

Data Fig. 1a shows the set of images used in our appearance manipulation study. They came from the three classes "chicken", "pear" and "fish". Subjects were asked to adjust the source images to become visually more similar to the respective target images. All participants had to edit the same set of images. The individual trials were shown in random order for each participant, where in each trial one image had to be edited using one of three methods:

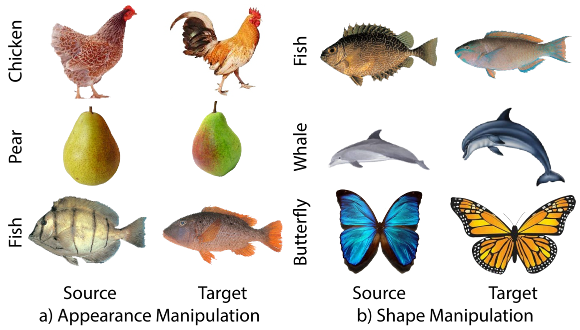

Figure 1: *The set of images used in our appearance* (a) *and shape* (b) *manipulation studies. We chose three different image pairs from three different classes for each manipulation task.*

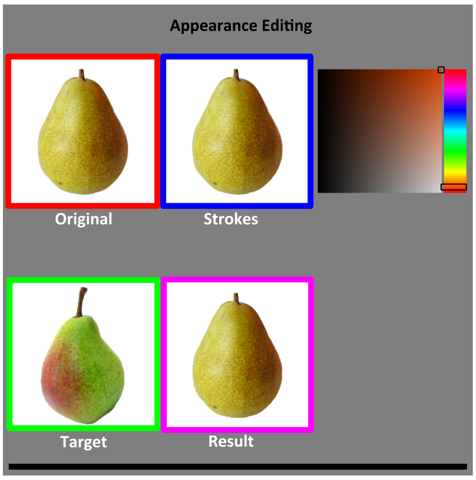

Figure 2: *A screenshot of the appearance manipulation study interface. The red, green, blue and pink rectangles (only for illustration in this figure) show the source and target image, the painting canvas and the output image, respectively.*

- Gastal and Oliviera's interactive colorization [GO11] using the user's paint strokes alone,
- Gastal and Oliviera's interactive colorization using our subspace as guidance instead of the strokes or
- our proposed subspace-guided locally-rigid appearance transfer.

Procedure At the beginning of the task, the following written instructions were presented to each subject:

> Thank you for participating in our study! Please indicate if you have any known difficulties in seeing colors before the experiment.
>
> Imagine you have a pear on the top left. Now you should adjust it to become visually more similar to the pear on

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.

the bottom left as quickly as possible. You can choose colors from the color picker and paint color strokes on the top right image and visualize the changes on the bottom right (See Fig.2).

You can press D to delete all of the current strokes and hold SHIFT while pressing the left mouse button to delete some parts of a stroke. You only have up to 20 seconds. After that, the next trial will be shown. The time left for each trial is shown in the black bar on the bottom. Adjustment is performed by choosing colors and painting strokes as often as you wish. The study will present 9 trials overall.

Next, a similar experiment will be performed with unlimited time. Please perform the manipulation as fast as you can. Please press ENTER to move to the next trial.

First, an initial training session will allow you to freely explore the principle of the study. We record no other data, than the color you select with every click.

3.2.3. Shape Manipulation Task

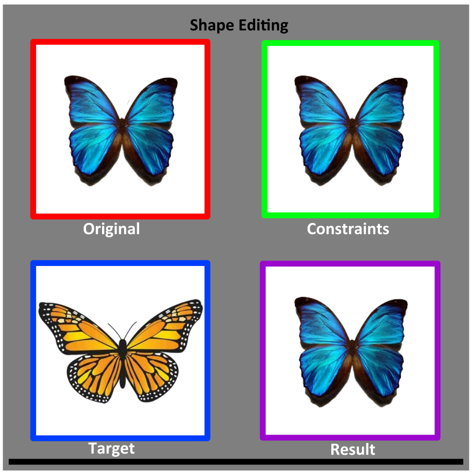

Figure 3: *A screenshot of the shape manipulation study interface. The red, blue, green and purple rectangles (only for illustration in this figure) show the source and target image, the constraints manipulation canvas and the output image, respectively.*

Data Fig. 1b shows the set of images used in the shape manipulation task. Again, we chose three different images from three classes: “fish”, “butterfly” and “whale”. The procedure was analogous to the one for the appearance manipulation task. Subjects were asked to adjust the source images to become visually more similar to the respective target images. The individual trials were shown in random order for each participant, where in each trial one image had to be edited either by a common manipulation method provided by means of Schaefer’s moving least squares (MLS) image deformation [SMW06] or our proposed subspace-aware shape manipulation.

Procedure At the beginning of the task, the following written instructions were presented to each subject:

Thank you for participating in our study.

Starting from the given butterfly on the top left, please deform it to become more similar to the fish on the bottom left as quickly as possible. The deformation can be done by putting position constraints on the top right image and visualize the changes on the bottom right (See Fig.3).

You can use left-mouse click to add more constraint points. The green circle shows the original position and red circle shows the new target position. You can drag the red circle around to change the target position. To delete a constraint pair, right click on the red or green circle. To delete all constraints, press D.

You only have up to 20 seconds. After that, the next trial will be shown. The time left for each trial is shown in the black bar on the bottom. Adjustment is performed by adding, deleting and modifying the position constraints as often as you wish. The study will present 6 trials overall.

Next, a similar experiment will be performed with unlimited time. Please perform the manipulation as fast as you. Please press ENTER to move to the next trial.

First, an initial training session will allow you to freely explore the principle of the study.

3.2.4. Manipulation Study Results

For both, appearance and shape manipulation, we performed two studies with limited and open-ended timing. Fig. 4 and Fig. 5 show the manipulated results achieved by one subject for the appearance and shape manipulation tasks, respectively.

In the appearance manipulation parts, we obtained $7 \times 3 = 21$ triplets of edited images from the seven subjects, for both, limited and open-ended time, where each triplet consists of two images edited using Gastal’s colorization with and without subspace and one image using our subspace-aware locally-rigid appearance transfer. Similarly, in the shape manipulation parts, we obtained $7 \times 3 = 21$ pairs, for both, limited and open-ended time, where each pair consists of one image edited using the reference and one using our subspace-guided method. In order to clarify whether our subspace improves the quality of manipulations, these results of the manipulation studies were used in a pairwise rating task, next.

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.

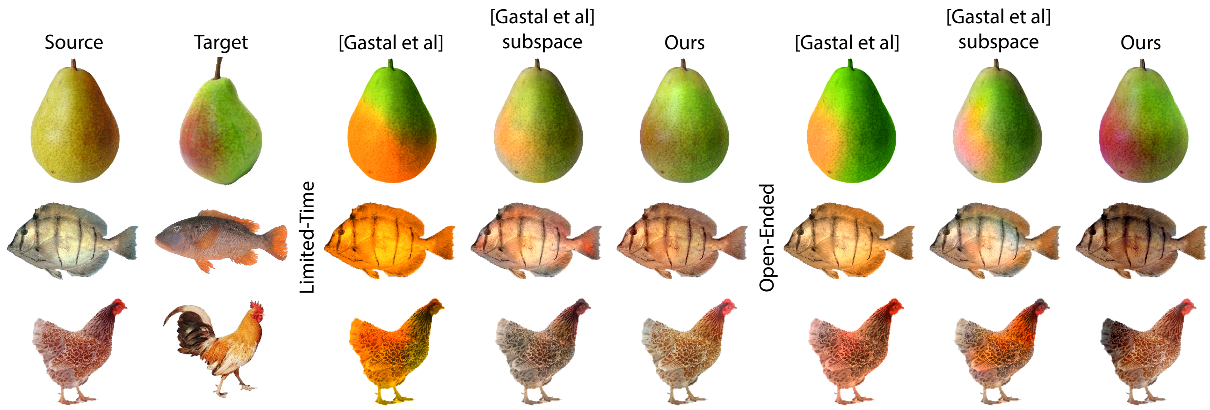

Figure 4: *Images manipulated by a subject for the appearance manipulation tasks.*

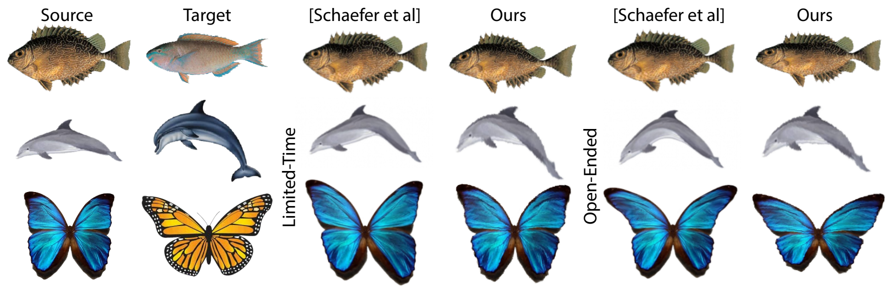

Figure 5: *Images manipulated by a subject for the shape manipulation tasks.*

3.3. Pairwise Rating Task

3.3.1. Participants

In this part of the study, we asked 12 subjects to rate the edited images obtained in the manipulation study (Sec. 3.2) in a pairwise rating task. None of them had participated in the previous study. Our subjects were between 23 and 50 years of age. Most of them were graduate students in engineering or art.

3.3.2. Data

To reduce the size of the pairwise rating task, from the appearance manipulation limited-time data we randomly picked 10 of the 21 triplets and generated a total of 30 pairs, where each pairs consisted of 2 out of the 3 methods that had been used to generate the images. Analogously, 30 pairs were picked from the appearance manipulation open-ended time data. From the shape manipulation data, we randomly picked 15 pairs (out of 21) for the limited time and 15 pairs for the open-ended setting. So, in total, we had $2 \times 30 + 2 \times 15 = 90$ pairs of images to be compared.

For each pair, we showed the target image next to pair of images generated by our and the reference methods, respectively, where the latter two were randomly permuted. Fig. 6 shows some "rows" shown to the subjects.

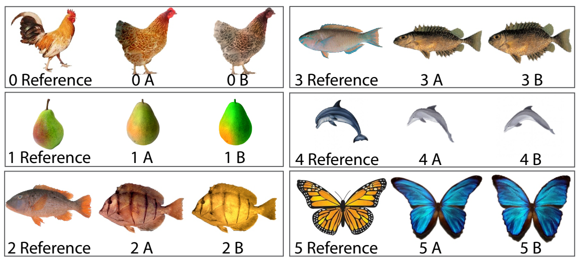

Figure 6: *Some pairwise comparison tasks shown to the subjects.*

Procedure At the beginning of the rating task, the following written instructions were presented to each subject:

> Thank you for participating in our study.
>
> In each row, the first image is given as a reference. Of the next two images, please choose the one that is visually closer to the reference.

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.

For example: 0) 0B

3.3.3. Pairwise Rating Results

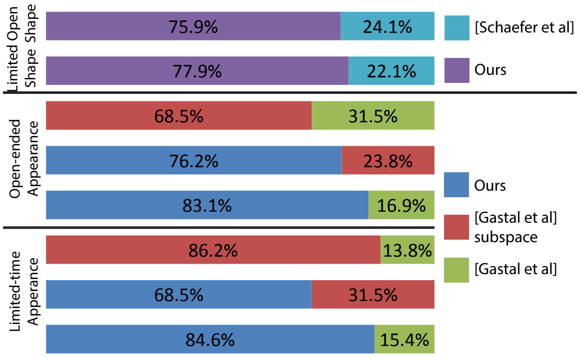

Figure 7: *Results of the pairwise rating tasks where each row shows the percentage of subjects that preferred either interface. The top two rows show the results for shape manipulation in the limited and open-ended settings. The bottom six rows show the results of the pairwise comparison tasks for the different appearance manipulation methods in the limited and open-ended settings.*

Fig. 7 shows the results of our pairwise rating task. All differences are statistically significant ($p < .01$, binomial test). For appearance manipulation, this indicates that our subspace-based locally-rigid transfer improves over the common colorization approach by Gastal [GO11] and that the latter could also be improved by using our subspace. Similarly, also our subspace-driven shape manipulation was preferred over simple MLS image deformation.

4. Image Manipulation

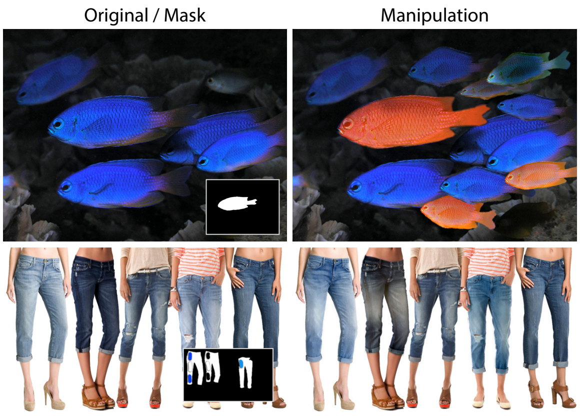

Figure 8: *After a user has marked an instance of the class* (inset)*, it can be used for cloning with new shape, new appearance or both from the space* (fish) *or to constrain color manipulation* (jeans).

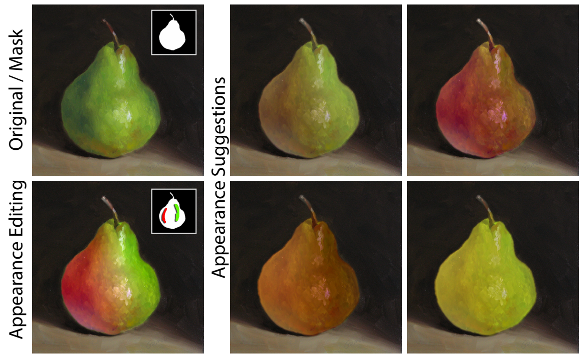

Figure 9: *Here, the user has marked an instance of the class pear* (inset) *in a still life painting. Our approach is used to constrain appearance manipulation and to suggest new appearances from the space of pears. Note how our appearance transfer method successfully preserves the style of the original painting.*

Figure 10: *From a given butterfly* (left, top)*, our system generate several shape suggetions. These are then used to stylize* (right) *the original image* (middle).

Fig. 8, Fig. 9, Fig. 10 and Fig. 11 show several complex images manipulation scenarios using our method. Fig. 12 demonstrates how our subspace can also be used in colorization of gray images.

5. Alignment

Fig. 13 shows additional alignment results for classes not shown in the paper. Fig. 16 shows the alignment graph of the boot dataset.

6. Transfer Matrix

Our method is based on factoring out appearance (Fig. 14, left), shape (Fig. 14, middle), or both (Fig. 14, right) to generate continuous re-combinations of the original images.

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.

Figure 12: *For an intial gray scale image* (a)*, a user provides a mask* (inset) *and paints green strokes to colorize the grass. For this part, Levin's colorization [LLW04] was used. The user then defines a mask* (inset) *for the chicken and paints some color strokes* (b)*.* (c) *shows the colorization results for Levin's colorization method.* (d) *shows how our subspace-aware appearance manipulation greatly helps the colorization process.* (e) *shows four suggestions computed based on the masked chicken* (b) *provided by user.*

Figure 11: *After a user has marked an instance of the class* (inset) *of an image* (top left)*, it can be used for appearance suggestions from the space of* horses (top right and bottom row).

7. Appearance Transfer

Fig. 15 shows different appearance transfer methods implemented either without or with support by our appearance subspaces. Fig. 17 shows how our locally-rigid appearance transfer behaves with different patch sizes.

References

[AP08] AN X., PELLACINI F.: Appprop: All-pairs appearance-space edit propagation. *ACM Trans. Graph. (Proc. SIGGRAPH) 27*, 3 (2008), 40:1–40:9. 8

[GO11] GASTAL E. S. L., OLIVEIRA M. M.: Domain transform for edge-aware image and video processing. *ACM Trans. Graph. (Proc. SIGGRAPH) 30*, 4 (July 2011), 69:1–69:12. 2, 5, 8

[HST13] HE K., SUN J., TANG X.: Guided image filtering. *IEEE PAMI 35*, 6 (2013), 1397–1409. 8

[LJH10] LI Y., JU T., HU S.-M.: Instant propagation of sparse edits on images and videos. *Comp. Graph. Forum (Proc. Pacific Graphics) 29*, 7 (2010), 2049–2054. 8

[LLW04] LEVIN A., LISCHINSKI D., WEISS Y.: Colorization using optimization. *ACM Trans. Graph. 23*, 3 (2004), 689–94. 6, 8

[SMW06] SCHAEFER S., MCPHAIL T., WARREN J.: Image deformation using moving least squares. *ACM Trans. Graph. (Proc. SIGGRAPH) 25*, 3 (2006), 533–40. 3

[XYJ13] XU L., YAN Q., JIA J.: A sparse control model for image and video editing. *ACM Trans. Graph. (Proc. SIGGRAPH Asia) 32*, 6 (2013), 197:1–197:10. 8

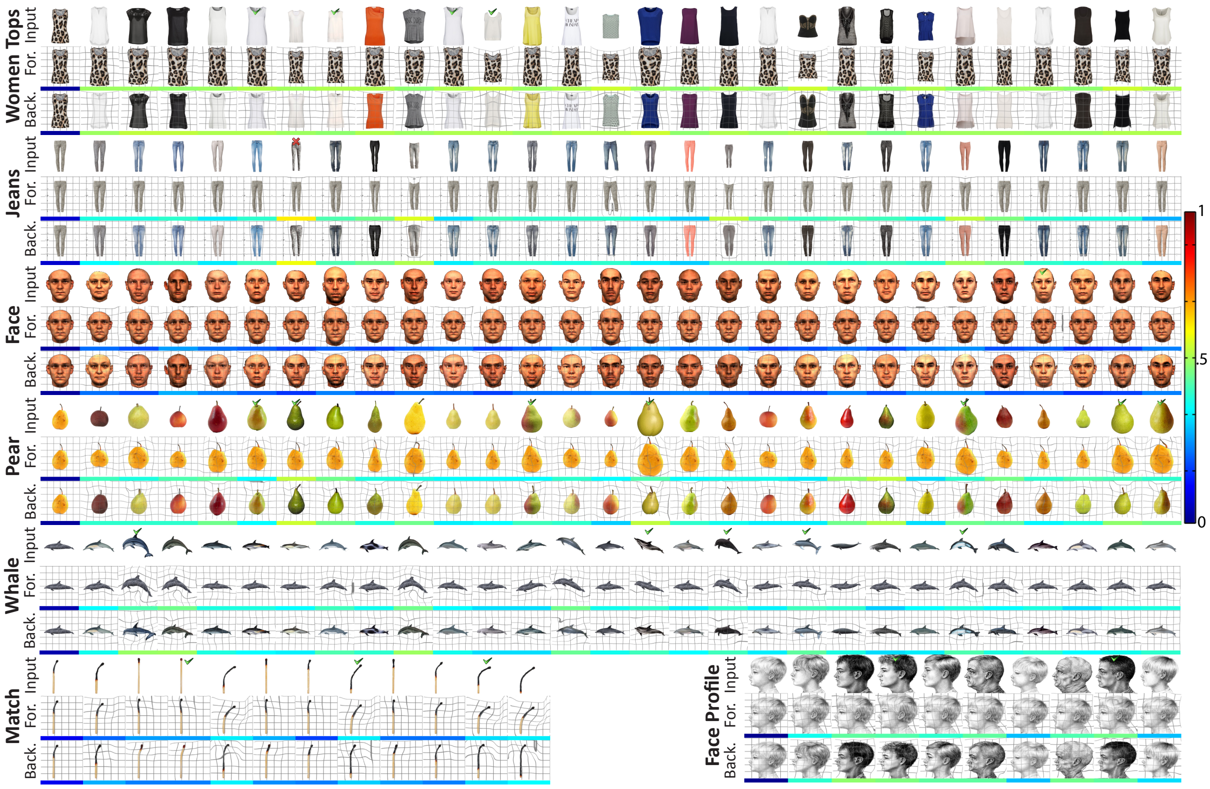

Figure 13: *Alignment of different exemplars* (columns) *from several classes* (rows). *For each class, the first column shows the reference image, the first row shows the input, the second row shows the forward alignment of the reference to the input and the third row shows the backward alignment of the input to the reference. The color below each exemplar indicates the match quality determined by SSIM, cf. the legend on the right. Smaller values indicate better alignment quality. Instances marked with a tick improved by using indirect alignment while instances marked with a cross were excluded because their image difference after alignment remained high.*

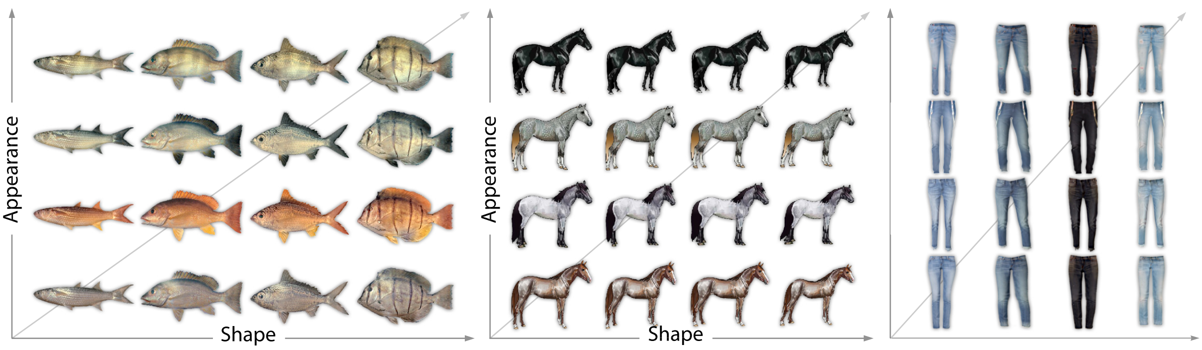

Figure 14: *We separate appearance of a collection of images (here shown on the diagonal) to create a space that contains arbitrary, continuous re-combinations (off-diagonal elements). For the* fish*, four differently shaped exemplars were combined with different appearances extracted from the input set. The middle figure shows four differently colored* horses *which in turn were combined with different shapes. Finally, for the matrix of* jeans *on the right, we varied both, appearance and shape, at the same time.*

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.

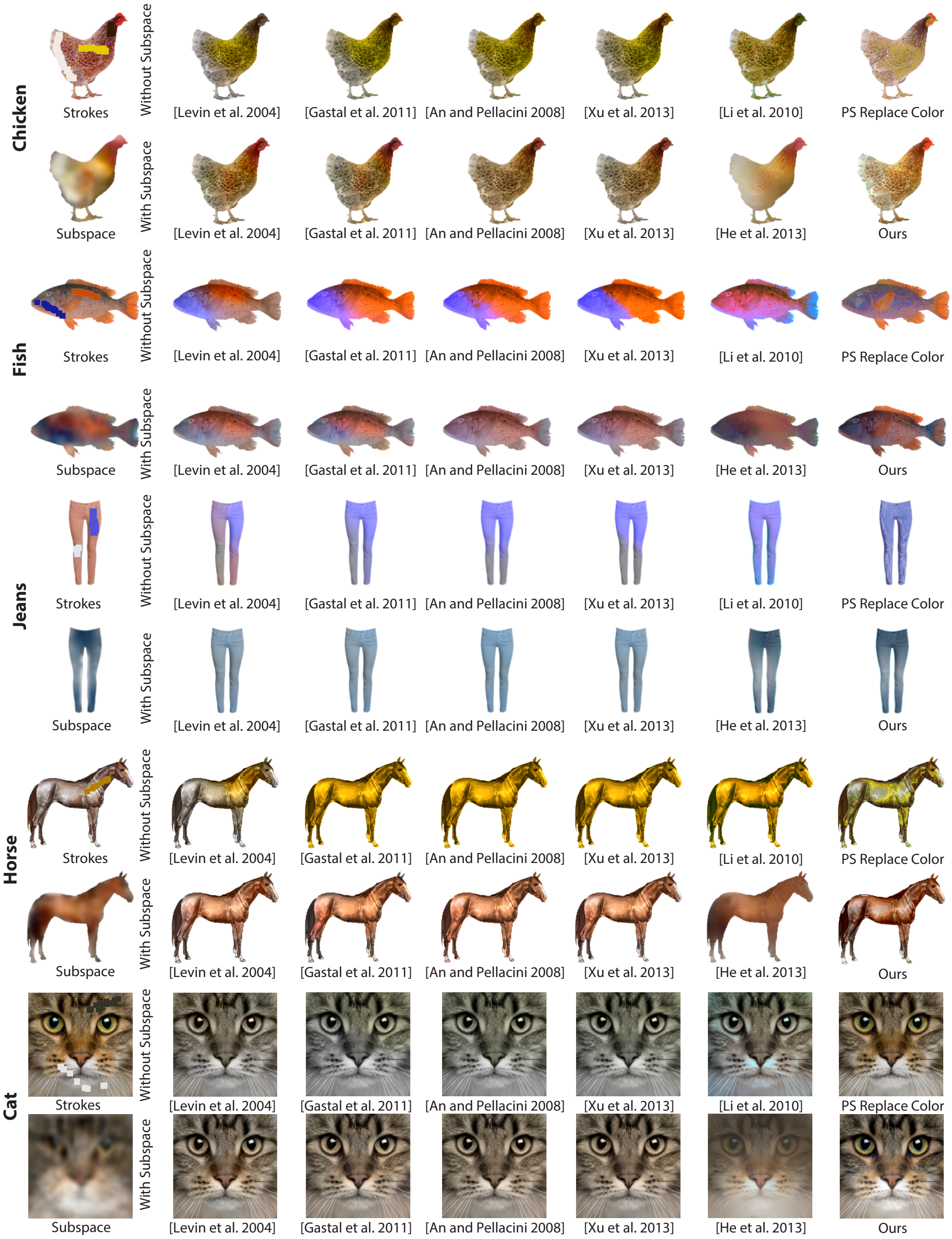

Figure 15: *Comparison of different appearance transfer methods. For each class, the first column shows the original with the user's input strokes* (top) *and the image reconstructed from the subspace* (bottom)*. Starting in the second column, the first row of every class shows stroke-based colorization using the methods of Levin et al. [LLW04], Gastal and Oliviera [GO11], An and Pellacini [AP08], Xu et al [XYJ13], Li et al. [LJH10] and Adobe Photoshop's "replace color" function, respectively. On the second row of every class, the second to the fifth column shows colorization given the subspace guidance using the methods used on the first row, the sixth column shows guided image filtering [HST13] using subspace as guidance and finally our locally-rigid appearance transfer on the final column.*

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.

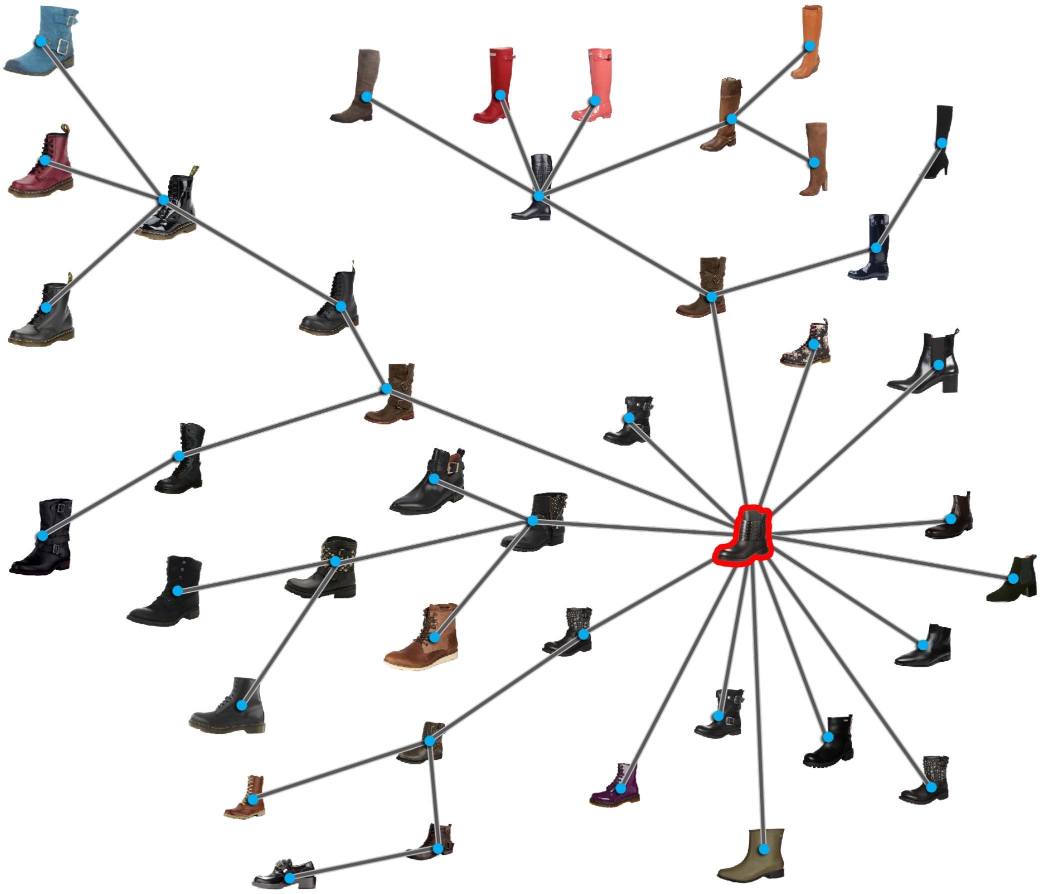

Figure 16: *The shortest-path tree of the "boot" class formed by the shortest paths from the reference (highlighted in red) to all other images in the class. Every node* (blue circle) *represents an image of the class and the undirected edges* (gray lines) *represent the connection between two nodes. An alignment path from the reference to another image is defined by walking through intermediate nodes using the edges in the graph.*

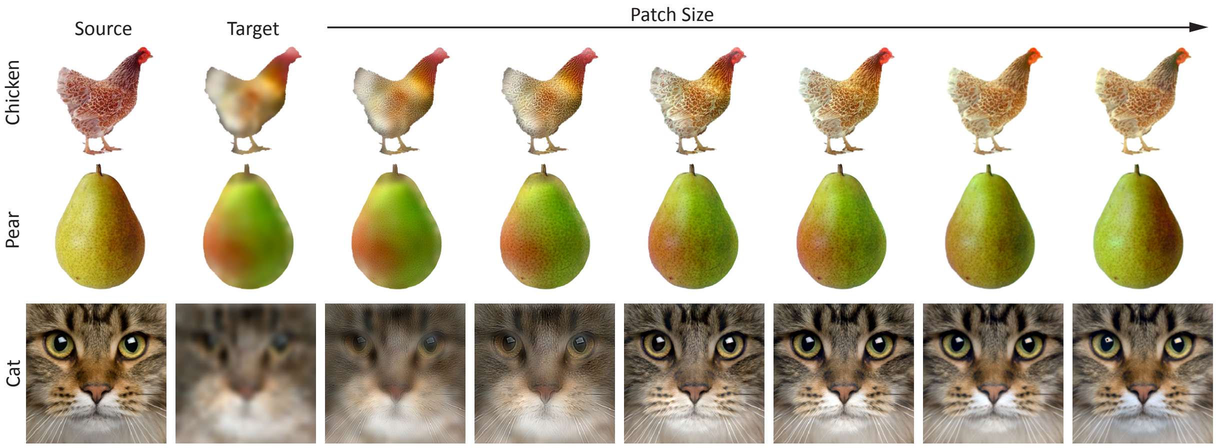

Figure 17: *The first and second column show the original and target images, respectively. From left to right are the results of our locally-rigid appearance transfer with increasing patch size.*

© 2015 The Author(s)
Computer Graphics Forum © 2015 The Eurographics Association and John Wiley & Sons Ltd.