Collage for Cover of PhotoBook

Marta Egorova, Ilia Safonov, Nikolay Korobkov
Moscow Engineering Physics Institute (State University), Russia
marfez_13 at inbox.ru, ilia.safonov at gmail.com; nikolay_korobkov at mail.ru

Abstract

We propose an algorithm for automatic creation of covers for photobooks in the form of collages. The appeal of different forms of collage from user's viewpoint is discussed. Both time and source-based selection that include estimation of the quality of the photograph are discussed. Image contrast, sharpness, and likelihood of presence of JPEG artefacts and color misbalance is assessed. The final step involves the arrangement of photos on canvases of various sizes by means of seamless blending. Obtained results for selection of photographs and constructed collages are evaluated using a survey. The majority of observers were pleased with the results of the algorithm.
Keywords: photo collage, photo quality estimation, photos clustering, photos arrangement, seamless blending.

1. INTRODUCTION

Recent years mark the development of the market for photo printing in the form of photobooks has begun to develop. Photobooks differ from traditional photoalbums in that all pages containing photos are printed jointly in the form of a bound book. One of the reasons that prevent wide use of photobooks is the time required to sort and lay out the photos. Automation of these processes is a topical theme. In this paper an algorithm of automatic creation of photobook covers in the form of collage is discussed. Such collage must:

- be part of photo selection process;
- appeal visually to a user;
- be constructed from photos that characterize the set of photos in a photobook in the best way.

It is assumed that a user makes only preliminary choice of photos; photos could be obtained from different cameras, have various sizes, quality and orientation. The number of photos in a typical photobook varies from tens to hundreds.

2. RELATED WORKS

There are many types of collages. Most frequently we see collages of 3 types: tiling or mosaic (Fig.1a) [1], photos heap up on table (Fig.1c) [2, 3], smooth transition between adjusted photos (Fig.1b) [4, 5]. It is worth noting the collage from paper [6], where the look of collage is chosen from predefined templates.

Automatic selection of representative photos from a given collection is an important part of collage creating algorithm. In [4] representative images are selected in three different ways: textually "interesting", mutually distinct and having faces present in the image. "Interestingness" of the image is assessed by computing entropy of histograms *ab* in color space *Lab*, mutual difference in distance between their histograms *ab*. Images with greatest entropy are considered as interesting. However, in our opinion it is not right to consider saliency from the viewpoint of the information theory. For example, it is known that image entropy increases with noise. Thus, images with a high level of noise may be selected for the collage. Assessment of similarity using distances in-between histograms of images can lead to images of different content being identified as similar.

Paper [4] lacks photo quality estimation, size and orientation are specified manually. In [1] blurred, underexposed and overexposed photos are excluded from analysis automatically. Time-based clustering is used to select photos for a slideshow. Algorithms for photo browsers use time-based clustering too [7, 8, 9, 10]. In addition these techniques try to exploit content-based information for selection, but their outcomes look ambiguous.

3. PLEASANT COLLAGE

In order to determine the most pleasant looking type of collage we conducted a survey. Survey participants were offered 5 selections of images with three different collages (a, b, c), as in Fig.1, as well as with varying number of images in collages of various selections (4, 6, 8, 15, 21). 30 participants were asked two questions: "Which collage is the most pleasant?" and "Which collage contains an excessive number of photos?" Answers to the first question are shown in Table 1.

Table 1 Results of survey about collage pleasantness

Image set/ Collage type	1	2	3	4	5	Total
a	5	3	12	12	9	41
b	18	20	13	16	20	87
c	7	7	5	2	1	22

Thus, the majority of observers prefer collages with smooth transition between photos, although the content of photos strongly affects answers. The majority thinks that using more than 8 photos for a collage for cover of a photobook is excessive. Taking these results into account, we decided to create a collage with *M* photos, where *M* ranges from 3 to 8 and can be adjusted by user. Alternatively, several collages are built for different values of *M* and the user then chooses the collage he likes most.

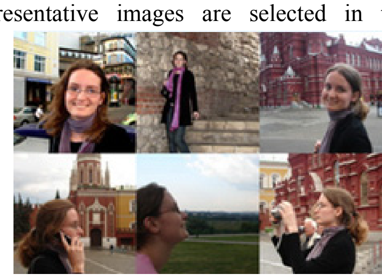
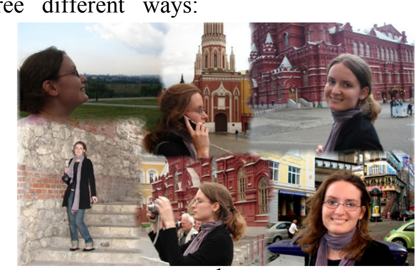
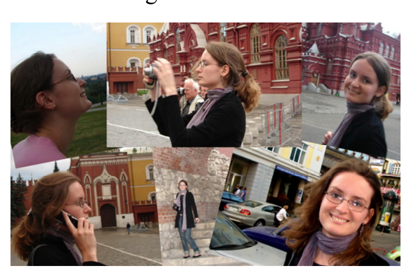

a

b

c

Figure 1 Collage comparison. Three ways to create a collage: (a) tiling collage, (b) collage with blending, (c) picture collage

4. REPRESENTATIVE IMAGE SELECTION

4.1. Estimation of photos quality

The first step in image selection for photobook is quality estimation. Photos with low quality are excluded from consideration. Quality is evaluated by looking at four different types of defects: blurriness, compression artefacts, low contrast, color misbalance. Paper [11] proposes an algorithm of non-reference sharpness estimation that is based on the idea that a high variation between the original and the blurred image means that the original image was sharp whereas a slight variation between the original and the blurred image means that the original image was already blurred. Sharpness metric A_{sub} is continuous magnitude and it has good correlation with subjective perception. We normalize sharpness estimation to the segment [0, 1]:

$$S = 1 - A_{sub} / 5 \;.$$

The metric for sharpness evaluation described in [12] can also be used. Such metric is also well correlated with subjective opinions. However, that metric is a four-grade discrete value, which is less convenient than a continuous sharpness estimate.

Global image contrast is determined from the brightness histogram:

$$C = (high - low)/255,$$

$$low = \min(\min\{i \mid H[i] \geq H_0\}, \min\{i \mid \sum_{k=0}^{i} H[k] \geq C_0\}),$$

$$high = \max(\max\{i \mid H_R[i] \geq H_1\}, \max\{i \mid \sum_{k=i}^{1} H_R[k] \geq C_1\}),$$

where H_0, H_1 and C_0, C_1 are threshold values for histogram area and intensity correspondingly.

JPEG compression artefacts reduce image quality. In [13] a filter for deblocking and deringing was proposed. For adjustment of filter parameters the top left corner of the square of 3x3 quantization table of brightness channel is analyzed:

$$K = \frac{1}{9} \sum_{i,j=1}^{3} q_{i,j}$$

Our research has shown that this metric correlates with visual assessment of JPEG images better than the compression ratio, which is strongly dependent upon content. In order to estimate the quality of JPEG images we used the following estimate:

$$A = \begin{cases} 0, & K > 30 \\ (30 - K)/30, & K \leq 30 \end{cases}.$$

For photos obtained from files of other format A=1. We exclude photos with low values S, C and A from the initial selection set of photos. The thresholds were chosen as a result of numerous experiments.

Sometimes photos are affected by color cast, for example, when indoor scene illuminated by an incandescent lamp is photographed. However, it is hard to distinguish between intentional color cast and cast caused by color misbalance. In our opinion, exclusion of photos with any dominating color from the photobook is unwise. However, exclusion of such photos from the cover is rational. In [14] various information on a huge data set of 4.8 million photos is analysed. Of those photos, 74% have no dominant colour that supports the general thesis about “gray world”. We do not consider photos for photo collage for which the mean of any channel differs from average brightness by more than 25%. Another important factor affecting image quality is noise. Unfortunately, we could not find an algorithm for noise estimation that would provide adequate results for real-world photos. We continue to research this area.

4.2. Time and source-based clustering

Time-based clustering provides satisfactory outcomes in terms of selection description in cases where all photos are photographed by one camera. When photos from different cameras are used, time-based clustering does not work as well.

We propose clusterization in 2D plane, where first axis is time and second axis is the source of photo. Here source is camera name obtained from EXIF information. If EXIF is absent, then all images for every present file format are merged into a separate source. Analysis in [14] paper shows that more than 70% of nowadays photos have EXIF tags. Fig.2 illustrates formulation rule for time-source 2D plane. Sources are sorted in ascending order according to the number of photos. L is time between the least and the most time for the largest source. Yp_i coordinate on Photo Source axis is calculated as follows:

$$Yp_i = \begin{cases} H \times (i+1)/2: & i \text{ is odd} \\ H \times (Nps + 1 - i/2): & i \text{ is even} \end{cases}$$

where i is index of source, Nps is number of sources, $H = L/M$.

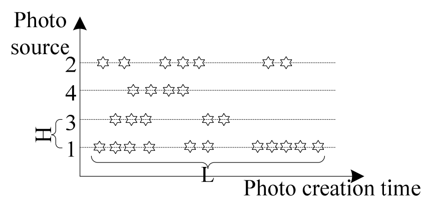


Figure 2 Time-source 2D plane

Next stage is clustering into M classes by applying *k-means* technique. From each class a photo with maximum integral quality estimation is selected for collage: $IQ_i = w_1 C_i + w_2 S_i + w_3 A_i$, where w_i are weights. If several photos have the same IQ, then the photo with maximum saturation is selected.

5. COLLAGE CREATION

5.1. Orientation detection

Initial images can be either portrait- or landscape-orientated. We determine the correct orientation and rotate the photo using information from EXIF. Unfortunately, part of the photos does not contain EXIF info or orientation tag. In this case it is advisable to use content-based algorithm of orientation determination [15].

5.2. Photo arrangement

Our approach allows not only construction of optimal layout of photos on a canvas, but also determination of scaling coefficients for every photo. In order to arrange the photos we first compose schemes of possible layouts. Divide canvas vertically and horizontally into M parts. Value of every cell ranges from 1 to M, which means in that part of canvas entry image with that number is located. For example, all possible ways to lay out 3 images are shown in Fig.3a.

A layout matrix for every identified way is constructed. We call layout matrix a matrix of 0 and 1, where 1 indicates that given image creates dimension w (width) or h (height).

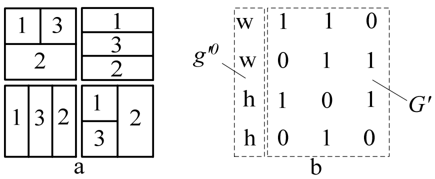


Figure 3 (a) schemes of possible layout of three images, (b) layout matrix bottom right plan

In order to determine coefficients of scaling we use least-squares method: $\overline{\alpha} = (G^T G)^{-1} G^T g^0$,

where $\overline{\alpha}$ is vector of scaling factors for each selected image;

$$\text{matrix } G_{ij} = \begin{cases} G'_{ij} w_i, & \text{if } \ g'^0_i = w, \\ G'_{ij} h_i, & \text{if } \ g'^0_i = h. \end{cases}$$

where w_i, h_i are width and height of *i-th* image, g'^0 and G' are created for each scheme;

$$\text{vector } g^0 \text{ is computed as: } g^0 = \begin{cases} w_c, & \text{if } \ g'^0_i = w \\ h_c, & \text{if } \ g'^0_i = h \end{cases}$$

where w_c, h_c are canvas width and height correspondingly.

Fig.3b shows g'^0 and G' for bottom right scheme from Fig.3a.

For example, canvas width is 800 and height is 600; selected images have sizes: (300, 400), (50, 150), (350, 150). For G' from Fig.3b values of G, g^0 and $\overline{\alpha}$ equal:

$$G = \begin{bmatrix} 300 & 50 & 0 \\ 0 & 50 & 350 \\ 400 & 0 & 350 \\ 0 & 150 & 0 \end{bmatrix}, \ g^0 = \begin{bmatrix} 800 \\ 800 \\ 600 \\ 600 \end{bmatrix}, \ \overline{\alpha} = \begin{bmatrix} 1.29 \\ 4.54 \\ 1.47 \end{bmatrix}$$

Discrepancy is computed as $(\psi(g_j) - g^0)$, where $\psi(g_j)$ is the value of resulting canvas size obtained from new sizes of images for given scheme. Using discrepancy, the best layout scheme is determined in terms of maximization of canvas area coverage and minimization of images overlap. Further scaling factors are changed by 15% in order to provide seamless smooth transition between adjacent photos.

5.3. Seamless blending

The scheme in Fig.4 describes the steps of our seamless blending approach with element of randomness.

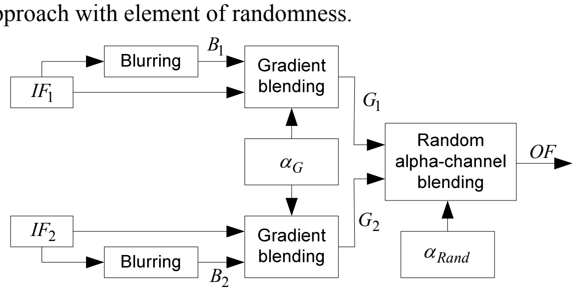


Figure 4 Steps of seamless blending approach

Suppose image I_1 and image I_2 are overlapped partially. Fragment IF_1 and fragment IF_2 are overlapped parts of image I_1 and image I_2 correspondingly. Fig.5 shows coordinate system: suppose the origin is left bottom angle of overlapped regions; axis x is directed horizontally and y vertically. Then x_{max} is the width and y_{max} is the height of the region. Our aim is to create a seamless transition in the region OF.

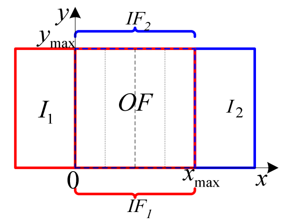


Figure 5 Coordinate system for overlapped regions

The first step is blurring overlapped parts of the images by applying FIR low-pass filter:

$$B_1 = h \otimes IF_1, \ B_2 = h \otimes IF_2$$

Blurred images compose with respective parts of the initial image:

$$G_1(x, y) = \alpha_G(x) B_1(x, y) + (1 - \alpha_G(x)) IF_1,$$

$$G_2(x, y) = (1 - \alpha_G(x)) B_2(x, y) + \alpha_G(x) IF_2,$$

where α_G specifies smooth function with random shape:

$$\alpha_G(x) = (2^{x^{(1+rand(2.5))}} - 1) / x_{max}.$$

Final step is blending of two photos via random alpha-channel α_{Rand}:

$$OF(x, y) = \alpha_{Rand}(x, y) G_1(x, y) + (1 - \alpha_{Rand}(x, y)) G_2(x, y),$$

$$\alpha_{Rand}(x, y) = \begin{cases} 1, & x \le T(x, y) \\ 0, & x > T(x, y) \end{cases},$$

where $T(x,y)$ is composite Catmull-Rom spline with the following control points P_i ($i = 0..N$):

$P_0 = (0, \ 0)$, $P_1 = (x_{max}/2, \ 0)$, $P_{N-1} = (x_{max}/2, \ y_{max})$,

$P_N = (0, \ y_{max})$, $P_{ix} = x_{max}/4 + rand(x_{max}/2)$, $P_{iy} = P_{(i-1)y} + C_{cp}$,

where $C_{cp} = y_{max}/30 + rand(y_{max}/30)$.

The final result of blending is shown in Fig.9. Our approach provides visually pleasant transition between photos.

6. RESULTS AND DISCUSSION

We conducted a survey to assess the results of photo selection process. Survey participants were asked to first look at the entire set of images and then look at two selected collections: one was done using the proposed approach and another using AutoCollage algorithm [4]. The question asked was: "Which one of the two offered collections characterizes the initial set of images better?"

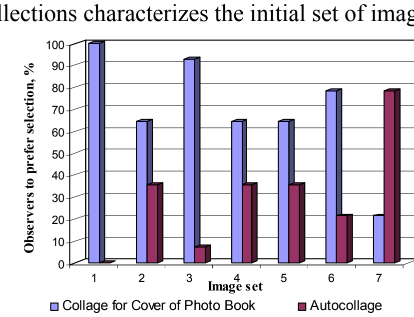


Figure 6 Survey results of representative collection

In total 7 collections containing between 40 and 205 images were offered. 14 people took part in the survey. Fig.6 presents results of our survey. It was noted for the majority of selections that the selection created using our method contains more typical and varying images. Selection created using AutoCollage contains brighter and saturated images. For example, for initial selection of images from the *ultimate frisbee* competition, shown in Fig.7, AutoCollage selected images shown in Fig.8.

Figure 7 Entire set of photos for selection

All selected images have saturated colors in comparison to photos of the actual game that took place in the sand. However, to insist that brighter images better characterize the collection of images is wrong. Also note that AutoCollage method has a tendency to select photos with color misbalance and often chooses blurred photos and photos of similar scenes.

Figure 8 Images selected by AutoCollage

In Fig.9 shown collage produced by the proposed method. The majority of observers were pleased with the created collages.

Figure 9 Collage produced by the proposed method

7. CONCLUSION AND FUTURE WORK

We propose method for creation collage cover of photobook which comprises selection of photos with quality estimation and arrangement with seamless blending. Demo application on C++ was developed too. Processing time for estimation of quality of one photo is about 1.5s; processing time for arrangement depends on number of photos in collage and varies in the range from 50 to 200s for notebook 1.9GHz and 256 Mb RAM. We plan to continue our investigations towards content-based selection and arrangement.

8. REFERENCES

1. Chen J., Chu W., Kuo J., Weng C., Wu J., "Tiling Slideshow: An Audiovisual Presentation Method for Consumer Photos", In Proc. of ACM Multimedia Conference, pp. 36-45, 2007.
2. J. Wang, J. Sun, L. Quan, X. Tang, H.-Y. Shum, "Picture Collage", In Proc. of CVPR06, pp.347-354, 2006.
3. S.Battiato, G.Ciocca, F.Gasparini, "Smart Photo Sticking", In Proc. 5^{th} Workshop on Adaptive Multimedia Retrieval, 2007.
4. C. Rother, L. Bordeaux, Y. Hamadi, A. Blake, "Autocollage", ACM Transaction on Graphics, vol 25, pp. 847-852, 2006.
5. T. Wang, T. Mei, X.-S. Hua, X. Liu, H.-Q. Zhou, "Video Collage: A Novel Presentation of Video Sequence", In Proc. of ICME, 2007.
6. B. Yang, T. Mei, L.-F. Sun, S.-Q. Yang, X.S. Hua, "Free-shaped Video Collage", Conf. on Multi-Media Modeling (MMM), pp. 175-185, Japan, 2008.
7. J.C. Platt, M. Czerwinski, B.A. Field, "PhotoTOC: Automatic Clustering for Browsing Personal Photographs", In *Proc.IEEE Pacific Rim Conf. Multimedia,* pp. 6-10, 2003.
8. J.C. Platt, "AutoAlbum: Clustering digital photographs using probabilistic model merging", In *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries,* pp. 96-100, 2000.
9. A. Graham, H. Garcia-Molina, A. Paepcke, T. Winograd, "Time as essence for photo browsing through personal digital libraries", Proc. of the 2^{nd} ACM/IEEE-CS joint conference on Digital libraries, USA, 2002.
10. D.F. Huynh, S.M. Drucker, P. Baudisch, C. Wong, "Time Quilt: Scaling up Zoomable Photo Browsers for Large, Unstructured Photo Collections", CHI'05 extended abstracts on Human factors in computing systems, USA, 2005.
11. F. Crete, T. Dolmiere, P.Ladret, M. Nicolas, "The Blur Effect: Perception and Estimation with a New No-Reference Perceptual Blur Metric", Proc. of SPIE/IS&T, v. 6492, 2007.
12. I.V. Safonov, M.N. Rychagov, K.M. Kang, S.H. Kim, "Adaptive sharpening of photos", Proc. of SPIE/IS&T, v. 6807, USA, 2008.
13. A. Foi, V. Katkovnik, K. Egiazarian. "Pointwise Shape-Adaptive DCT for High-Quality Denoising and Deblocking of Grayscale and Color Images". *IEEE Transaction on Image Processing*, v.16, No 5, May 2007.
14. D.Wueller, R.Fageth,"Statistic Analysis of Millions of Digital Photos", Proc. of SPIE/IS&T, v. 6817, USA, 2008.
15. E.Tolstaya, "Content-based image orientation recognition", GRAPHICON-2007, pp. 158-161, Russia, 2007.

About the authors

Marta Egorova received her MS degree in cybernetics from Moscow Engineering Physics Institute/State University (MEPhI), Russia in 2008. At present time she is post-graduate student of MEPhI.

Ilia Safonov received his MS degree in automatic and electronic engineering from MEPhI, Russia in 1994 and his PhD degree in computer science from MEPhI in 1997.

Nikolay Korobkov received his MS degree in cybernetics from MEPhI in 1988 and his PhD degree in computer science from MEPhI in 1997.