석 사 학 위 논 문
Master's Thesis

이미지 검색을 활용한 자가촬영 사진의 배경 합성

Scene Completion of Selfie Photos based on Image Search

2016

윤 웅 직 (尹 雄 稷 Youn, WoongJick)

한 국 과 학 기 술 원

Korea Advanced Institute of Science and Technology

석 사 학 위 논 문

이미지 검색을 활용한 자가촬영 사진의 배경 합성

2016

윤 웅 직

한 국 과 학 기 술 원

전산학부 (웹사이언스대학원)

이미지 검색을 활용한 자가촬영 사진의 배경 합성

윤 웅 직

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회의 심사를 통과하였음

2016년 6월 16일

심사위원장 윤 성 의 (인)

심 사 위 원 김 동 준 (인)

심 사 위 원 최 기 선 (인)

Scene Completion of Selfie Photos based on Image Search

WoongJick Youn

Advisor: Sung-Eui Yoon

A dissertation submitted to the faculty of
Korea Advanced Institute of Science and Technology in
partial fulfillment of the requirements for the degree of
Master of Science in Engineering in Computing (Web Science Technology)

Daejeon, Korea
June 16, 2016

Approved by

Sung-Eui Yoon
Professor of Computing

The study was conducted in accordance with Code of Research Ethics[1].

[1] Declaration of Ethical Conduct in Research: I, as a graduate student of Korea Advanced Institute of Science and Technology, hereby declare that I have not committed any act that may damage the credibility of my research. This includes, but is not limited to, falsification, thesis written by someone else, distortion of research findings, and plagiarism. I confirm that my thesis contains honest conclusions based on my own careful research under the guidance of my thesis advisor.

MWST 20143437

윤웅직. 이미지 검색을 활용한 자가촬영 사진의 배경 합성. 전산학부 (웹사이언스대학원). 2016년. ??+ii 쪽. 지도교수: 윤성의. (영문 논문)
WoongJick Youn. Scene Completion of Selfie Photos based on Image Search. School of Computing (Graduate School of Web Science Technology). 2016. ??+ii pages. Advisor: Sung-Eui Yoon. (Text in English)

초 록

본 논문에서는 기존의 풍경 사진 합성방식과 다르게 백만 장의 이미지 데이터셋을 기반으로 자가촬영 사진의 주변 배경을 확장해 나가며 합성하는 알고리즘을 제안하였다. 추출한 이미지 기술자의 효과적인 검색을 위하여 근사 근접 이웃 기반의 프러덕트 양자화 알고리즘을 이용해 기술자를 양자화하여 유사한 사진을 검색하였고, 사진 속 기술자의 위치 정보를 활용하기 위해 블록 형식의 격자 형태로 나누어서 기술자에 가중치를 주었다. 이러한 방식은 색상뿐만 아니라 사진의 질감 또한 고려하여 원본 사진과 유사한 사진의 검색이 가능하고, 유사한 사진에 포아송 이미지 에디팅 기법을 사용하여 반복적으로 주변 사진을 합성하여 결과물을 생성할 수 있다. 이를 통하여 영상 정보가 존재하지 않는 사진 밖 주변 영역으로 사진을 확장시켜 나갈 수 있음을 확인하였다. 또한, 사용자는 기존 원본 사진보다 몇 배 확장 할지만 결정하면 기존의 사진 합성 방식처럼 합성할 영역을 특정하지 않아도 새로운 사진을 생성하는 장점이 있다.

핵 심 낱 말 영상 합성, 이미지 검색, 근사 근접 이웃 검색

Abstract

We address scene completion of selfie photos through image search technique to expand the background of those photos. For searching descriptors efficiently, our system quantizes image descriptors for approximate nearest neighbor search, while finding similar images. Moreover, a query image is divided into a grid for weighting to utilize the spatial information of those descriptors. It allows searching similar images based on color and texture so that a result image can be completed with pasting similar images on the border of the query image by Poisson image editing iteratively. Therefore, our system can extrapolate the selfie photo by expanding the background regions where visual information does not exist. After determining the size of result image, our system also generates a creative image without specifying regions for image completion.

Keywords Image completion, Image search, Approximate nearest neighbor search

Contents

List of Figures

Chapter 1. INTRODUCTION

Creating a synthetic image has been widely studied in image processing and computer graphics. People can generate the synthetic image through many computing tools, but it is not easy to create a visually plausible image. Therefore, there are many studies to improve the quality of the synthetic result by an image completion, and these can be categorized into two families by a source for filling.

Classical techniques for the image completion are intended for restoring damaged photographs and removing unwanted regions in an image. The unwanted regions are looked like a hole so a visual information from surrounded areas can be utilized to fill the regions. Partial differential equations based [2] and examplar-based algorithms [3, 13] are used to fill the hole with the information from the source image, and a structure information is also utilized in recent studies [8, 17]. However, filling the regions with the visual information from the source image is inevitable from having the duplicated content.

On the other hand, a result based on a data-driven approach can be completed with similar images from an image collection. Hays and Efros [5] proposed the novel algorithm of a scene completion utilizing millions of images. The nearest neighbor search helps to find similar images to create a synthetic image without a complicated formula, but a texture mismatch can be led because it does not handle a texture information. In a recent research, graph matching is used for scene completion [19]. Multiple levels of graphs are extracted from hierarchically segmented regions, and similar images are searched based on these graphs for an extrapolation. However, such methods restrict the class of image library for concentrating on a related content and keeping the size of image collection which can prevent the system becoming slow.

To solve this issue, the approximate nearest neighbor search has been researched for an efficient searching in a large-scale dataset. Recently, product quantization [11] showed high accuracy by decomposing the vector into disjoint subvectors and using codewords instead of exact data points. Since this method does not use all data points while calculating the distance, product quantization only requires loading codewords on the main memory, so it improves memory efficiency and the cost for searching.

In this paper, we propose a novel scene completion algorithm for generating an expanded selfie photograph with a proper background. The goal of this task is to create a larger selfie photo by filling the boundary regions based on a data-driven approach. Our system quantizes image descriptors from a large image collection with product quantization to generate codebooks. After codebook is generated, a target region which includes a query image and marginal regions to be filled is divided into patches. Each patch region searches a similar image through image search to be completed respectively. We use the optimized version of Poisson image editing [18] on only boundary regions of each patch to preserve the original information of two images so it can help suppressing drastic color changes. Finally, the result image is generated by iterating this sequence until all the patches are completed.

To demonstrate benefits of our method, we test our method on MIRFlickr-1M dataset and compare our result over content-aware fill tool in Photoshop. Content-aware fill performs well on monotonous patterns for generating a seamless result, but our result contains a creative content on background regions thanks to a data-driven approach. Moreover, our system prevents the duplication problem of a person from the selfie photo because our system does not utilize the visual information for the photo in a copy and paste manner.

Chapter 2. RELATED WORK and BACKGROUND

In this section, we review techniques of classical image completion, and prior works of scenery image completion and approximate nearest neighbor search.

2.1 Image Completion

Image completion has been widely used to remove an unwanted portion from images and to restore the damaged parts of old photographs. This technique fills user-specified regions by utilizing the information of surrounding areas to obtain a visually acceptable image. Bertalmio et al. [2] proposed a texture synthesis algorithm based on partial differential equations (PDE). However, example-based image completion algorithms [3, 13] have been shown more realistic inpainting results in natural images compared to PDE-based algorithms.

Example-based methods synthesize the texture of missing regions from small patches in a stochastic manner to generate a large pattern for filling. Example-based algorithms also utilize structure information to handle structure inconsistency. Sun et al. [17] used user-specific curves to extend its structures from known regions. Huang et al. [8] suggested a patch-based approach using mid-level structural information without any user input.

These methods perform well for generating synthetic images such as hole-filling. Nonetheless, there is a limitation that unknown regions are only filled with the information from the source image.

2.2 Scene Completion

Hays and Efros [5] proposed a scene completion algorithm in a data-driven approach by utilizing millions of Flickr images. Through nearest neighbor search, the algorithm finds a few hundreds of candidate images that are similar to the source image and generates synthetic images with graph-cut and Poisson blending [16]. Thanks to this data-driven method, visually plausible results can be completed without a complicated formula of classical image completion approach. Since candidate images are only searched by utilizing GIST [15] descriptors, a texture information is not handled properly during scene matching. As a result, it can lead to a texture mismatch, resulting in sub-optimal results.

Recently, scene completion also incorporates graph matching for search process [19]. After a user specifies directions to be extrapolated, images are hierarchically segmented to build graphs in multiple levels. Each node denotes a region in the source image, while each edge represents the relationship of adjacent regions. Once graphs are constructed with local features in segmented regions, subgraph matching is used near the boundary of extrapolation direction in a coarse-to-fine matching.

Our approach is based on the scene completion method to extrapolate a selfie photo with product quantization, enabling scalable image search with one million of images.

2.3 Approximate Nearest Neighbor search

Approximate Nearest Neighbor (ANN) search has been studied on compact encoding in a high dimensional vector space. Among ANN search techniques, hashing [6, 10, 20] has received attention, since it can generate compact vectors whose similarities are computed by the Hamming distance. Recently, Product Quantization (PQ) techniques [7, 11] decompose the original vector space into disjoint subspaces by k-means clustering, and the distance between two data points is calculated based on their codewords.

PQ have been improved in many different directions thanks to its high accuracy. Ge et al. presented Optimized Product Quantization (OPQ) [4] to minimize the quantization distortions with space decomposition and optimal codebooks. Kalantidis and Avrithis [12] improved OPQ further to optimize codes per segments locally to encode residuals. Unlike previous PQ techniques, Babenko and Lempitsky proposed Additive Quantization (AQ) [1] that encodes a vector using the sum of codewords from different dictionaries. Zhang et al. introduced Composite Quantization [21] that also represents a vector by the summation of inner products in a similar manner to AQ.

We use PQ in our method to efficiently perform image search. Since other improved PQ techniques can be used straightforwardly used in the place of PQ, those techniques can be used within our method.

Chapter 3. OUR APPROACH

In this section, we describe our approach for extending selfie photos with extracted features through image search. We first explain the dataset and features what we have extracted (Sec.3.2). We then show the approach of image search to find candidate images based on product quantization (Sec.3.3). Lastly, we describe the process of extending the background of selfie photos (Sec.3.4).

3.1 Overview

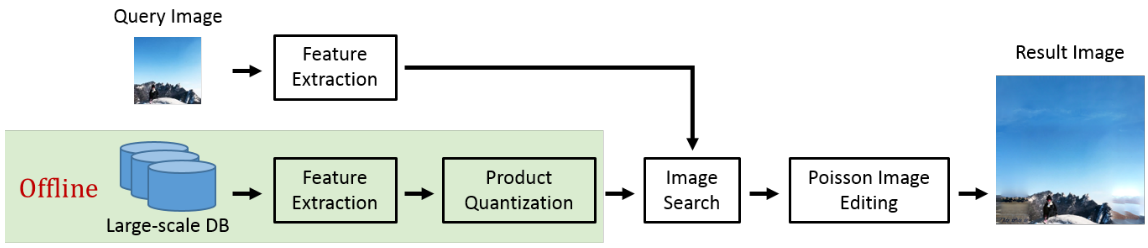


Figure 3.1: The overview of our approach. In a green region, we extract descriptors and generate codebooks from product quantization in the offline stage. In online stage, the descriptors of a query image are extracted and search similar images from the codebooks. Then, the similar images are completed with the query image with Poisson image editing technique for generating a result.

Extrapolating the image is a challenging work compared to other image completion techniques that use hole-filling algorithms. During the process of hole-filling, the visual information from surrounding areas of a source image can give a cue for missing regions. The naive method for extrapolating a source image is to identify similar patches near the boundary of the source image. However, identifying similar patches to those boundary regions within the source image is extremely difficult. This is mainly because regions which are needed to be extrapolated hold a significant portion of a result image. Therefore, those regions need to be filled with similar but creative contents. For this reason, an image collection that has various sorts of photographs is essential for our problem, because the result will be affected by images from the collection that will be pasted on unknown areas.

On the other hand, nearest neighbor search in high-dimensional data is not scalable to deal with finding similar images. A large-scale image collection with high-dimensional data demands a high-performance machine with sufficient memories to load, which means that it is restricted to run on typical computers. Searching based on approximate nearest neighbor techniques overcomes the limitation by encoding data with preserving its representation, and it allows efficient use of memory.

In this paper, we handle a scene completion for a selfie photograph to extend the background. Our goal is to generate an enlarged selfie by filling the boundary regions by a data-driven approach. Since we use the selfie photo as an input and assume that a person is located in the bottom center of a query, our system expands boundaries in three directions (left, right, and up) so that we can prevent the body of the person from locating in the center of the result. Through the data-driven approach, our system finds a candidate image from a collection instead of exploiting the visual information within the query image.

The overview of our system is illustrated in Fig. 3.1. We first extract descriptors from a large-scale image collection, then generate codebooks of those descriptors by Product Quantization (PQ). We create a result image which is background expanded from the query, and the size of result image is decided by a user input. The result image including extended regions is divided into patches, and each patch region searches a similar photograph for image completion through PQ respectively. Only boundary regions in the patch are completed by a Poisson image editing method to preserve the original information of two images for merging instead of completing the whole patch. It prevents changing the color of a source image by a target image and fills a background more plausible while preserving the visual data of the query. We generate the result by iterating this process until all the patches are completed.

3.2 Dataset and Feature Extraction

We need various kinds of pictures to demonstrate our system, so we tested our system on MIRFlickr-1M dataset [9] which has one million images from Flickr. In previous works [5, 19], they only used an image collection of outdoor scenes for reasonable search time. However, we do not restrict classes or tags of images in the dataset showing a robustness to diverse photographs which can include unrelated to the scene photograph.

We use GIST descriptors which can depict a picture in generally, and it performs well on searching similar images based on known regions of a query. Extracting a GIST descriptor requires a color transformation from RGB to LAB color space, but gray-scale images only have a light information in LAB color space. Therefore, we exclude gray-scale images for the consistency of dimensions. We extract gist descriptors in eight by eight spatial bins with four scales for comparing overlapped blocks of source and target images, so the dimension of GIST descriptor is 4,608. Since gist descriptors do not handle texture well, we also extract Local Binary Patterns (LBP) [14] descriptors from HSV color space to consider a texture information. As a result, three channels of LBP descriptors are extracted as HSV color space has three visual properties. LBP descriptors consist of eight by eight spatial bins to utilize blocks in a similar manner to GIST, and the dimension of LBP descriptor is 3,776 on each HSV channel. We dealt only two types of descriptor but other global descriptors, e.g. CNN and HOG, also can be used for our system.

3.3 Image Search based on Product Quantization

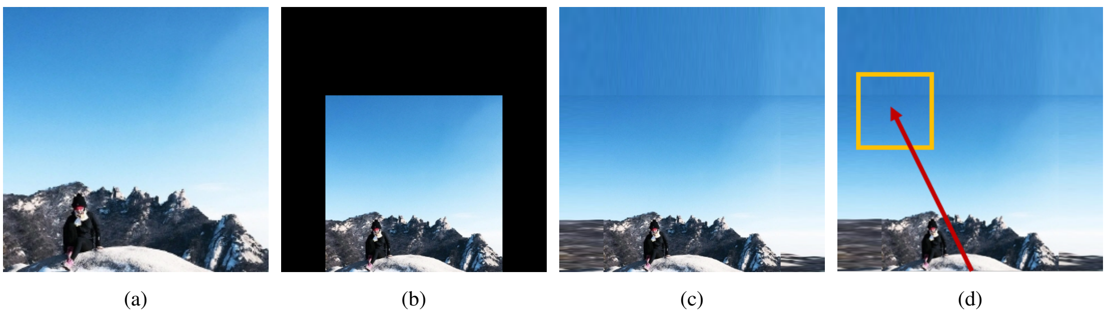


Figure 3.2: (a) is a query image and the black region of (b) is a target region of the result. (c) is the expansion of marginal regions. (d) shows the patch-based image search.

We use patch-based image search to fill expanded background of result image. After a user specifies the size of result image, the result is comprised of a query image and marginal regions where to be filled later. Marginal regions are filled with a small portion of query's boundaries to have a cue for the background. Since these marginal regions are used as a source for searching and blend with candidate images, the result image can preserve its color consistency. The result is then divided into small patches for a search and completion. The process of this task is illustrated in Fig. 3.2.

Nearest neighbor search requires loading all the descriptors on the main memory, so it leads memory inefficiency while searching similar images for every patch. Therefore, our system uses product quantization to reduce search cost. We generate codebooks of descriptors from an image dataset by PQ. As we noted in 3.2, GIST and LBP-HSV descriptors are comprised with eight by eight spatial bins, so we divide each descriptor into 64 subspaces for PQ, and each subspace is quantized to 256 codewords by k-means clustering. The codebook of PQ assigns the descriptors from a query to similar images in a collection. During the assigning process, we use asymmetric distance computation which calculates the Euclidean distance between the query and centroids of each subspace as shown in fig.3.1.

Let us denote vector x as a query and vector y as an image collection. These vectors are concatenated with m subvectors, and each subvector is quantized to codewords c_i which are centroids of codebook C. We define the Euclidean distance $\| x-y \|$ as d(x,y) in following equations. For asymmetric distance computation, we quantize the vector y to the nearest codeword of i^{th} codebook as follows:

$$q_i(x) = arg\min_{c_i} d(x,c_i) \tag{3.1}$$

The distance of gist descriptors is defined as:

$$d_{gist}(x,y) = \sqrt{\sum_{i=1}^{m} d\big(x,q_i(y)\big)^2} \tag{3.2}$$

For LBP descriptors, we calculate each visual property of HSV color space respectively:

$$d_h(x,y) = \sqrt{\sum_{i=1}^{m} d\big(x,q_i(y)\big)^2} \tag{3.3}$$

$$d_s(x,y) = \sqrt{\sum_{i=1}^{m} d\big(x,q_i(y)\big)^2} \tag{3.4}$$

$$d_v(x,y) = \sqrt{\sum_{i=1}^{m} d\big(x,q_i(y)\big)^2} \tag{3.5}$$

As a result, the distance between the query and image collection is computed as:

$$dist(x,y) = d_{gist}(x,y) + d_h(x,y) + d_s(x,y) + d_v(x,y) \tag{3.6}$$

After the distance computation, we pick the nearest vector from the image collection for a scene completion. Our system iterates Sec. 3.3 and Sec. 3.4 to fill all the patches of result image.

3.4 Scene Completion

As we find a target image which is similar to source image in a patch, we use Poisson image technique to blend both images within the patch. Poisson image editing [16] performs well on completing two images in seamless cloning, but it has a limitation that the color of the target image is not preserved. Therefore, we use the optimized algorithm of Poisson image editing [18] which can control the level of color adaptation with a color preserving parameter. This parameter suppresses drastic color changes in the background of the result, so our system can avoid propagating dissimilar colors of the source image. Moreover, our system only pastes the boundaries of the result to the patch while preserving the visual data of both images as shown in Fig.3.3

(a) A target image (b) A source image from dataset (c) The result after image completion

Figure 3.3: (a) is a target image and red region is for blending. (b) is a source image for image editing and the only red region will be completed with a target region.(c) is the result of blending two images. The Red region is completed only with the source image, and the visual information on rest of the region is preserved.

Chapter 4. Experimental Results

We have implemented our system with Matlab and compared our result with content-aware fill tool in Photoshop CC. For our experiments, we extracted and quantized the descriptors on a machine with 3.47GHz Intel Xeon CPU and 72GB RAM. On the other hand, we ran our system on a computer with 3.4GHz Intel Core i7 CPU and 32GB RAM as our system does not require a high-performance machine.

4.1 Dataset

The goal of our system is to generate a larger selfie photo by pasting similar images near boundaries. We avoid the duplication of a person in the selfie photo with searching visually plausible photographs from an image dataset. Therefore, our system requires a large-scale image collection to exploit the benefit of a data-driven approach. For the image collection, we do not restrict classes or tags to show the robustness of dependency on a dataset. We tested our algorithm with one million Flickr images as noted in 3.2. If we use a 256 by 256 sized query image, it takes about 13 seconds to generate 1.5 times larger result. Even though the dataset contains unrelated photographs, our system finds similar images for background in a reasonable search time.

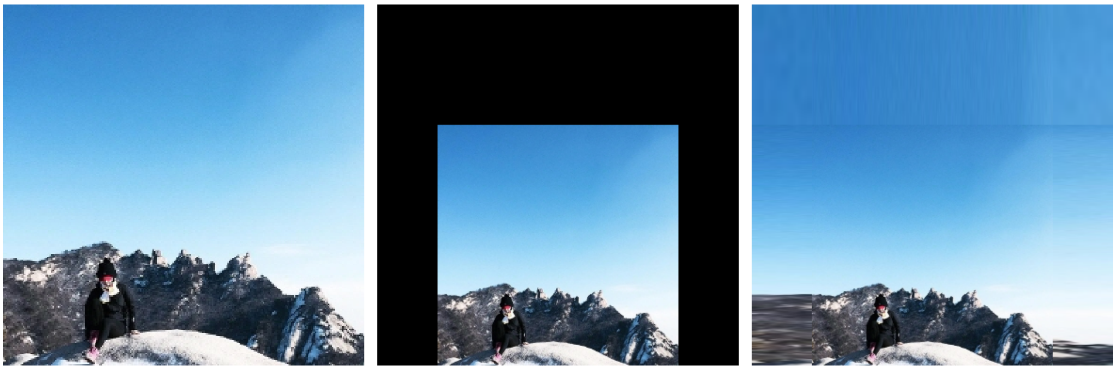
GIST descriptors have 4,608 dimensions for eight by eight spatial bins with four scales, and each LBP descriptors of HSV have 3,776 dimensions in eight by eight spatial bins. As a result, the size of GIST descriptors is about 32GB, and each LBP descriptor of HSV has about 26GB, which is about 110GB in total for both descriptors. However, product quantization allowed compressing descriptors with 256 codewords, which is 8 bits, in disjoint 64 subspaces. Therefore, the size of GIST and LBP descriptors are compressed into about 260MB. This memory load is about 450 times smaller than original descriptors. Our system searches in one million images by using a small amount of memory, so we do not need a high-performance machine to run our system.

4.2 Patch-based Image Expansion

Our system divides an expanded query image into a grid for patch-based image expansion. Each patch is completed with a similar image which has the smallest sum of distance (Eq. 3.6). However, our system needs to sort the order of patches for image completion by picking the closer patches first. In this order, we expand patches from the boundaries of query image rather than choosing further patches, so the image completion is propagated from the query region naturally. Moreover, the result of a patch is improved by referring the previous result where the region is overlapped with previous image completion. It can generate and extend the object, e.g. a cloud or a sky, which is not in the query image but is created by surrounding patches. As you can see in Fig. 4.1, our system generates a cloud while expanding the background region of the sky.

4.3 Scene Completion Result

Since our approach is different from previous methods, it is difficult to compare directly with their results. Previous works only find one similar image to complete for a specific region but our system combines some similar images by pasting on background regions to complete. Therefore, we compare our result with content-aware fill



(a) A query image (b) The target region of result (c) The expansion of marginal regions

(d) (e) (f) (g)

(h) (i) (j) (k)

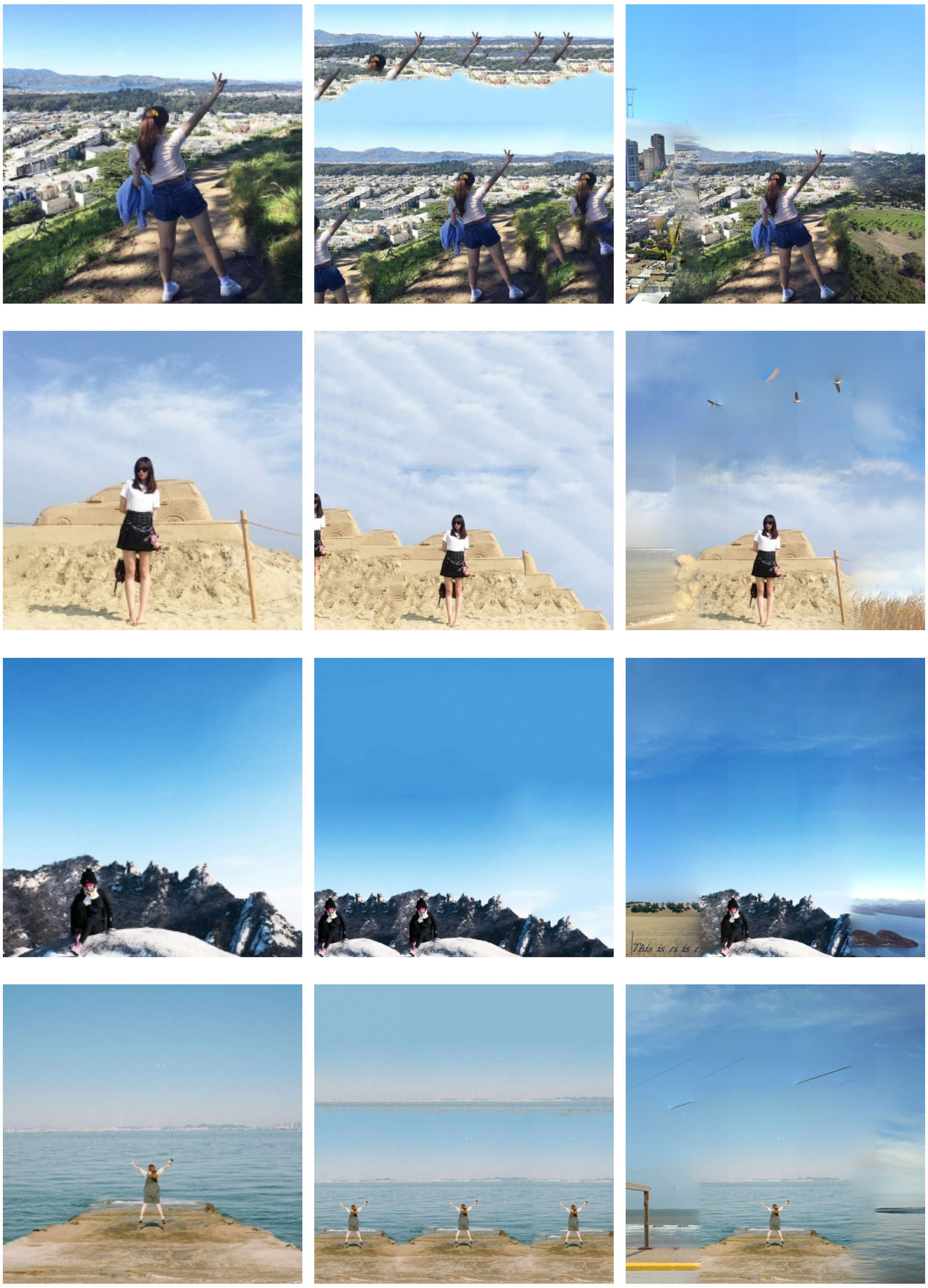
(l) (m) (n) (o)

Figure 4.1: (a) is a query image for scene completion and the black region of (b) represents a area to be filled. (c) is the expansion of marginal regions where to be completed. The sequence of (d) to (o) shows a step for patch based image completion. A cloud does not exist over the sky in the query image, but some white clouds are created in (o).

tool in Photoshop CC. We tested selfie photo as a query which has a scenery background, and a result is generated in 1.5 times larger than the query image for a comparison.

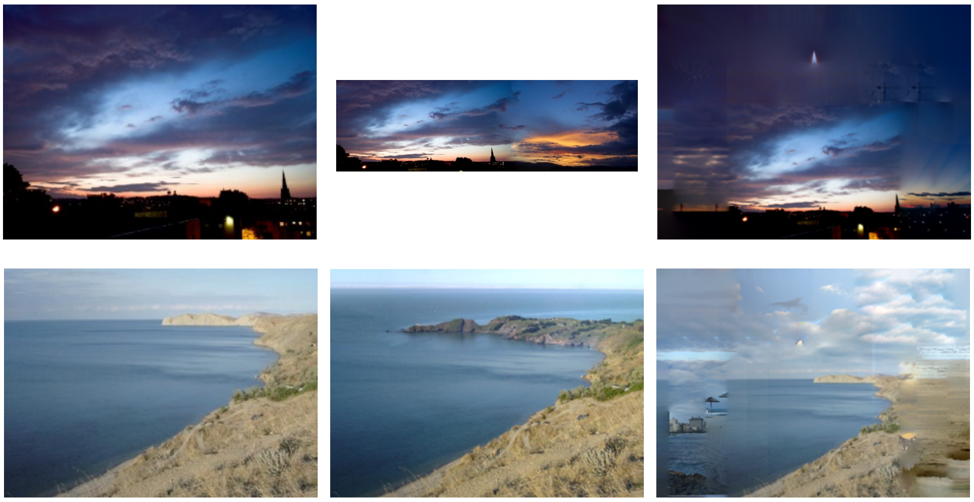
As you can see in Fig. 4.2, content-aware fill utilizes the information of the inner region to fill target regions in a copy and paste manner. It performs well on monotonous patterns for extrapolating a seamless result, but a person in a selfie photo can be duplicated in the result. Our result prevents the duplication problem because our system does not utilize the visual information of the query image. We also compare our result with the previous extrapolation algorithm [19] in 4.3. Wang et al.'s algorithm shows a natural result with one photograph for scene completion, but our results contain dissimilar patches. However, our system generates each result in only 13 seconds with low memory footprint compared with several minutes for the segmentation process for the previous work. Moreover, our results generate some creative contents on the background regions which do not exist in the query, and this usually happens in the sky regions of results as shown in Fig. 4.4.

Other the other hand, some of our results contain odd contents that are from the image collection. As descriptors are divided into eight by eight spatial bins for searching, similar images which include odd contents can not be filtered properly during the search process, and it can be solved by using dense spatial bins for PQ. Moreover, previous works performed well based on millions of photographs with restricting tags for image libraries. A data-driven approach depends on the amount of image collection and the quality of them. Therefore, our system also can be improved by exploiting larger dataset and utilizing image tags which are provided.



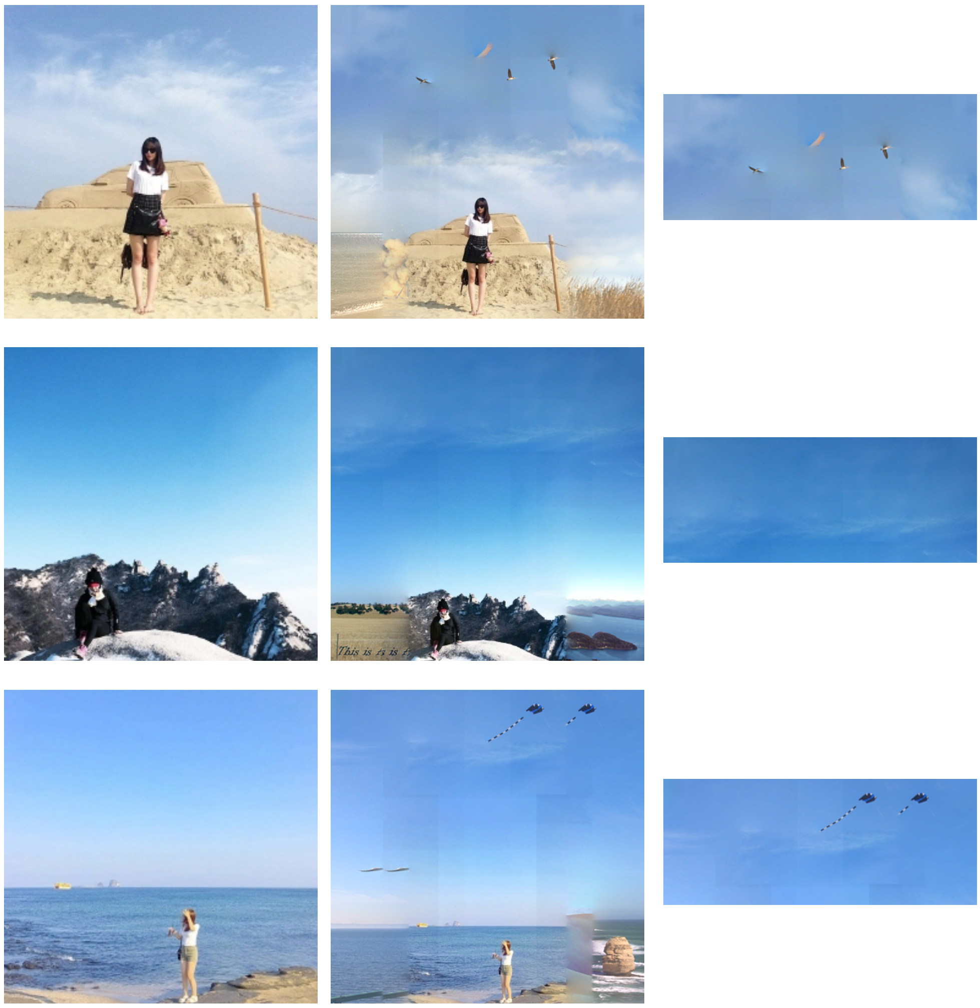
(a) Query images (b) Content-aware fill (c) Ours

Figure 4.2: (a) images are query for scene completion and (b) images are the result of content-aware fill tool in Photoshop CC. (c) images are the result of our system.



(a) Query images (b) Wang et al. (c) Ours

Figure 4.3: (a) images are queries and (b) images are the result of Wang et al.'s algorithm. (c) images are the result of our system.



(a) Query images (b) Ours (c) Magnified images

Figure 4.4: (a) images are the query and (b) images are the result of our system. (c) images represent magnified images of the skies in (b), and a creative content can be found in our results.

Chapter 5. CONCLUSION & FUTURE WORKS

We have proposed a novel scene completion algorithm for extrapolating the background of a selfie photo. We utilize a data-driven approach for image search, so our system generates a visually plausible image without a complicated formula. Therefore, our system uses patch-based image completion to expand the boundary of the selfie photo and quantizes image descriptors by an approximate nearest neighbor technique. Moreover, each patch is divided into a grid to utilize the spatial information of global descriptors. Through this approach, our method searches similar images based on color and texture, and a result image is completed by pasting those images with Poisson image editing iteratively. Finally, our system generates a larger selfie photo with extrapolating the background where visual information does not exist. We test our system on one million Flickr images and compare our result with content-aware fill tool in Photoshop CC. Since our system does not utilize the visual information of the query image in a copy and paste manner, our result prevents the duplication of a person from the selfie photo. Moreover, our result contains a creative content on the background regions thanks to the data-driven approach.

As a result image from the data-driven approach depends on the size of data, our system generates a more plausible image with a larger image collection. We can also utilize a tag to categorize image dataset which is related to a scenery photograph. On the other hand, the result image can be completed naturally by improving search process. In our experiment, our system used eight by eight spatial bins but searching performance will be improved by using more spatial bins. Moreover, the recent version of PQ methods can be applied to our system for enhancing the quality of the result.

Bibliography

[1] A. Babenko and V. Lempitsky. Additive quantization for extreme vector compression. In *CVPR*, pages 931–938, 2014.

[2] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image inpainting. In *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 417–424. ACM Press/Addison-Wesley Publishing Co., 2000.

[3] A. Criminisi, P. Pérez, and K. Toyama. Region filling and object removal by exemplar-based image inpainting. *TIP*, 13(9):1200–1212, 2004.

[4] T. Ge, K. He, Q. Ke, and J. Sun. Optimized product quantization for approximate nearest neighbor search. In *CVPR*, 2013.

[5] J. Hays and A. A. Efros. Scene completion using millions of photographs. *TOG*, 26(3):4, 2007.

[6] J.-P. Heo, Y. Lee, J. He, S.-F. Chang, and S.-E. Yoon. Spherical hashing. In *CVPR*, 2012.

[7] J.-P. Heo, Z. Lin, and S.-E. Yoon. Distance encoded product quantization. In *CVPR*, 2014.

[8] J.-B. Huang, S. B. Kang, N. Ahuja, and J. Kopf. Image completion using planar structure guidance. *TOG*, 33(4):129, 2014.

[9] M. J. Huiskes, B. Thomee, and M. S. Lew. New trends and ideas in visual concept detection: the mir flickr retrieval evaluation initiative. In *Proceedings of the international conference on Multimedia information retrieval*, pages 527–536. ACM, 2010.

[10] P. Indyk and R. Motwani. Approximate nearest neighbors: toward removing the curse of dimensionality. In *STOC*, 1998.

[11] H. Jégou, M. Douze, and C. Schmid. Product quantization for nearest neighbor search. *IEEE TPAMI*, 2011.

[12] Y. Kalantidis and Y. Avrithis. Locally optimized product quantization for approximate nearest neighbor search. In *CVPR*, 2014.

[13] V. Kwatra, A. Schödl, I. Essa, G. Turk, and A. Bobick. Graphcut textures: image and video synthesis using graph cuts. In *TOG*, volume 22, pages 277–286. ACM, 2003.

[14] T. Ojala, M. Pietikäinen, and T. Mäenpää. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE TPAMI*, 24(7):971–987, 2002.

[15] A. Oliva and A. Torralba. Modeling the shape of the scene: a holistic representation of the spatial envelope. *IJCV*, 2001.

[16] P. Pérez, M. Gangnet, and A. Blake. Poisson image editing. In *TOG*, volume 22, pages 313–318. ACM, 2003.

[17] J. Sun, L. Yuan, J. Jia, and H.-Y. Shum. Image completion with structure propagation. In *TOG*, volume 24, pages 861–868. ACM, 2005.

[18] M. Tanaka, R. Kamio, and M. Okutomi. Seamless image cloning by a closed form solution of a modified poisson problem. In *SIGGRAPH Asia 2012 Posters*, page 15. ACM, 2012.

[19] M. Wang, Y.-K. Lai, Y. Liang, R. R. Martin, and S.-M. Hu. Biggerpicture: data-driven image extrapolation using graph matching. *TOG*, 33(6):173, 2014.

[20] Y. Weiss, A. Torralba, and R. Fergus. Spectral hashing. In *NIPS*, 2008.

[21] T. Zhang, C. Du, and J. Wang. Composite quantization for approximate nearest neighbor search. In *ICML*, pages 838–846, 2014.

Acknowledgments in Korean

Curriculum Vitae in Korean

이　　　름: 윤 웅 직
생 년 월 일: 1986년 10월 04일
주　　　소: 대전 유성구 대학로 291 한국과학기술원 전산학부 3443호
전 자 주 소: woongjick.youn@kaist.ac.kr

학　　력

2002. 3. – 2005. 2.　서울영동고등학교

2006. 9. – 2010. 5.　Univ. of Minnesota Twin Cities Computer Science (B.S.) Mathematics Minor

경　　력

2010. 12. – 2013. 12.　(주)오토브레인 연구원