http://dx.doi.org/10.5392/JKCA.2012.12.03.067

천 커버링의 원리와 알고리즘 그리고 언어 식별에 응용

Principle and Algorithm of Cloth Covering and Application to Script Identification

김민우*, 오일석**
전북대학교 전자정보공학부 컴퓨터공학*, 전북대학교 컴퓨터공학부/영상정보신기술연구소**

Minwoo Kim(bghunter@chonbuk.ac.kr)*, Il-Seok Oh(isoh@chonbuk.ac.kr)**

요약

본 논문은 천 커버링 연산의 개념과 알고리즘을 제안한다. 천 커버링은 물리 법칙에 기반을 둔 연산으로 사물을 덮는 천의 모양을 계산학적으로 흉내낸다. 천 커버링의 목적은 사물을 천으로 덮어 표면의 상세함을 감추고 개략적인 외형이 드러나도록 하는 것이다. 이 연산은 천의 뻣뻣한 정도를 제어하는 하나의 크기 인자를 가지며, 이를 통해 외부로 드러나는 사물에 대한 정보의 상세함을 조절한다. 제안하는 연산의 가능성을 보이기 위해 문서 영상에 사용된 언어를 식별하는 문제에 천 커버링을 적용하였다. 실험 결과 가우시안을 이용한 특징 추출 방법보다 천 커버링을 이용한 특징 추출 방법이 더 우수한 식별 성능을 보였다. 토론에서 제안하는 연산이 우수한 이유를 제시한다.

■ 중심어 : | 다중 규모 | 특징 추출 | 언어 식별 | 신호 처리 |

Abstract

This paper proposes a concept and algorithm of cloth covering. It is a physically-based model which simulates computationally a shape of cloth covering some objects. The goal of cloth covering is to conceal the details of object and to reveal only the shape outline. It has one scale parameter which controls the degree of suppressing fine-scale structures. To show viability of the proposed cloth covering, this paper performed an experiment of script recognition. The results of comparing accuracies of feature extraction using Gaussian and cloth covering showed that the cloth covering is superior to Gaussian. We discuss the reason for the superiority.

■ keyword : | Multi-scale | Feature Extraction | Script Identification | Signal Processing |

I. Introduction

The multi-scale representation of signals or images is important issue in various research communities such as signal processing, computer vision, and pattern recognition [1]. With the purpose of multi-scale representation, this paper proposes a novel model which computationally simulates the physical phenomenon of cloth covering. The initial cloth on ground level is lifted under predefined rules to cover the objects until it reaches a stable state. In this process, the cloth is modeled as a list of balls

* 이 논문은 2011년도 정부(교육과학기술부)의 재원으로 한국연구재단의 기초연구사업 지원을 받아 수행된 것임 (2010-0010737)
접수번호 : #120116-003
접수일자 : 2012년 01월 16일
심사완료일 : 2012년 02월 03일
교신저자 : 오일석, e-mail : isoh@jbnu.ac.kr

connected by rubber band.

Since the objects covered by the cloth lose their details and appear only in their outlines, basic effect of the cloth covering operation is suppressing fine-scale structures of given signals or images. In other words, the operation transforms microscopic structure into macroscopic one. It has one scale parameter which controls the degree of suppressing fine-scale structures.

To show viability of the proposed cloth covering, this paper performed an experiment of script recognition. Text line images written in three different scripts were used by the experiment. Coefficients taken from Fourier transforms of profile and projection of line images were used as features. For feature extraction, Gaussian and cloth covering were applied to the raw profile and projection signals. The results of comparing their accuracies showed that the cloth covering is superior to Gaussian.

In the Discussion section, we attempt to explain why the proposed cloth covering is superior to conventional method. The essential property is related to the structure-preserving nature of the cloth covering. This property is considered to be the source of higher discriminative of the proposed features.

Section 2 presents the related works and Section 3 presents the principle and algorithm of the cloth covering. Section 4 describes experiments done using script recognition problem. Section 5 concludes the paper.

II. Related Works

The scale-space theory supported by multi-scale representation methods is motivated from biological vision and physics. The most popular method is using Gaussian kernel as a smoothing operator [2]. The Gaussian has one scale parameter which controls the degree of suppressing fine-scale structures in given signals or images. The scale-space theory has many applications which could benefit from scale-invariance property, such as feature extraction (e.g., SIFT [3]), edge detection [4][5], and interest feature point detection [6].

Another line of approach is wavelet theory which represents signals using orthonomal set of basis functions [7]. However this method does not generate an explicit shape of the input signals. Actually the Gaussian is the most popularly used method for suppressing fine-scale details.

In computer graphics, a similar approach called LOD (Level of Detail) has been studied actively [8]. It involves decreasing the complexity of 3D objects with the aim of faster rendering. When an object is far from the camera viewpoint, it is coarsely rendered since the human eyes cannot recognize the details of the object. An actual application of LOD approach can be found in cloth simuation [9]. However, the aim of computer graphics is different from the signal processing. In computer graphics, the aim is to reduce the computational loads for a realtime rendering. In signal processing as this paper, the aim is to obtain a more appropriate signal for the specific application area. For example, in our application of the script identification of Section IV, the aim is to obtain the more discriminative feature set.

III. Principles and Algorithms

As [Figure 1] shows, the cloth is modeled as a list of cloth elements connected by elastic material. The elastic material is an object with elasticity such as rubber band or spring. This paper calls it as *rubber band*. The cloth element is an object with weight.

This paper calls it as *ball*. The paper assumes all the balls have the same weight and the rubber band is weightless. It also assumes that the length of rubber band is proportional to the weight given to the rubber band, i.e., proportional to the number of balls hung from the rubber band.

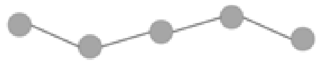


Figure 1. The cloth is modeled by a list of balls connected by rubber band.

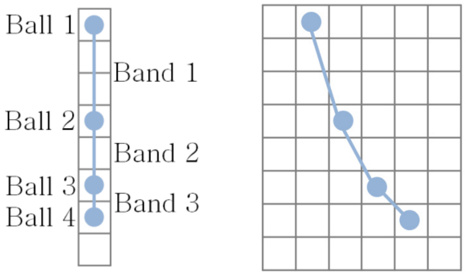
1. Dynamics

Assume that two balls with weight m are connected by a rubber band with elasticity coefficient k. When a ball is lifted up high enough that the other ball is hung in the air, the resulting length of rubber band (denoted by x) is proportional to the force gived on the rubber band. We can formulate $f=kx$ by Hook's law. Since the only force on the rubber band is the weight of hanging ball, we can write $f=mg$ where g is the gravitational acceleration. So the length of rubber band is $x=mg/k$ [10].

Let's extend the situation to one where four balls are connected as in [Figure 2](a). In this situation, the length of rubber band 3 is equal to mg/k. However since two balls are hung from rubber band 2, the length of rubber band 2 is $2mg/k$. Similarly the length of rubber band 1 is $3mg/k$.

Let's generalize the phenomenon. When n balls are hung from a rubber band, the length of the rubber band is nmg/k. Since m, g, k are constant, let's replace $k/(mg)$ by τ. Now the length of rubber band from which n balls are hung can be written by Equation (1).

$$x = \frac{1}{\tau}n \qquad (1)$$



(a) Balls at the same column (b) Balls at neighbor columns

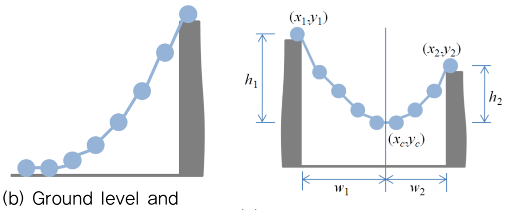
Figure 2. Four balls hung from top

In Equation (1), τ is the parameter which controls stiffness of the cloth. The larger the τ is, the stiffer the cloth is. On the contrary, as τ becomes smaller, the cloth becomes smooth and the final shape is similar to the original signal. We call τ as *stiffness coefficient*.

In order to define the cloth covering operation over 1-dimensional signal, let's change the situation of [Figure 2](a) to [Figure 2](b). In the new situation, four balls are not located on the same column but on adjacent columns. Though physically the length of rubber band should be calculated considering vertical and horizontal forces, this paper considers only the vertical direction for computational simplicity. So Equation (1) can be applied to the situation of [Figure 2](b).

[Figure 3] illustrates three primitive situations which algorithm of the cloth covering should deal with. [Figure 3](a) is the case where the signal is ground level. In this case, the cloth is on ground level. [Figure 3](b) is one where one side is ground level and the other side is on top of supporting bar. In this case, the cloth shape can be calculated using Equation (1). [Figure 3](c) shows the case where both sides are on top of supporting bars. Since in this case the cloth can't be calculated simply by using Equation (1), we divide the rubber band into two parts. We assume that two parts are independent, i.e., they do not influence on the resulting shape each other. We take the lowest ball as the point where the rubber band is split into two parts. After splitting the rubber band, Equation (1) can be applied to each of two parts.

(a) Ground level



(b) Ground level and supporting bar

(c) Two supporting bars

Figure 3. Three primitive situations

Now we will explain the procedure to find out the lowest ball for the case of [Figure 3](c). Let the lowest point to be (x_c, y_c) and the coordinates of top of supporting bars to be (x_1, y_1) and (x_2, y_2). Then the height can be written as $h_1 = y_1 - y_c$ and $h_2 = y_2 - y_c$. And the width are $w_1 = x_c - x_1$ and $w_2 = x_2 - x_c$. The values of h_1 and h_2 can be calculated by summing Equation (1) as shown in Equation (2).

$$h_1 = \frac{1}{\tau}\sum_{i=1}^{w_1} i, \quad h_2 = \frac{1}{\tau}\sum_{i=1}^{w_2} i \qquad (2)$$

Since the left and right parts are connected at the lowest point, the equation, $y_c = y_1 - h_1 = y_2 - h_2$ holds. Rearranging the equation, we get Equation (3).

$$x_c = \frac{2\tau(y_1 - y_2) - (x_1^2 - x_2^2 - x_1 - x_2)}{-2(x_1 - x_2 - 1)} \qquad (3)$$

2. Algorithm

In the situations of [Figures 3], balls at the ends are fixed on ground or on top of supporting bars. In other words, we know two contact points where the ball is contact with signal level. However, in the initial state, no contact point is known. This section explains the algorithm which starts with the initial state and ends with the final cloth shape.

Algorithm 1. 1-dimensional cloth covering

```
Input: D[0,n-1]    //input signal
Output: C[0, n-1] // cloth shape

1. C[i] = 0, 0≤i≤n-1
2. Covering(0, n-1);


Covering(p, q)
{
   if(p == q) return;

   pivot = argmax_{p≤k≤q} (D[k] - C[k]);

   if(D[pivot] - C[pivot] ≤ 0) return;
   C[pivot] = D[pivot];

   Physics(p, pivot);
   Covering(p, pivot);

   Physics(pivot, q);
   Covering(pivot, q);
}

Physics(p, q)
{
    update cloth shape in the range [p,q] using
  dynamics in Section 3.1;
}
```

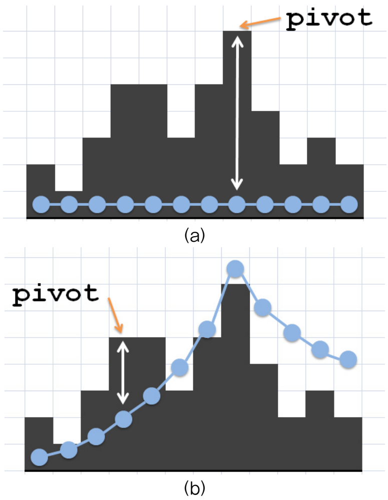


Figure 4. Algorithm for cloth covering (recursively splitting at pivots)

[Figure 4](a) shows the initial state which the algorithm starts with. In this state, the algorithm attempts to decide a ball which is doomed to contact with cloth. The point is called as pivot point. The criterion to choose the pivot point is maximizing the height difference between current cloth and the input signal. The pivot point is indicated in [Figure 4](a).

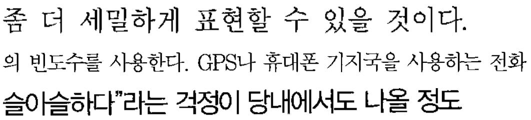
The cloth is lifted at the pivot so that the cloth is contact with the signal level at the pivot as [Figure 4](b) shows. Then the cloth is split into two parts at the pivot, and two parts proceed the same procedure independently. [Figure 4](b) indicates the new pivot for the left part.

The algorithm can be written in a recursive procedure in Algorithm 1. The algorithm starts with ground level cloth of C[i]=0, i=0,n−1. The recursive function Covering(p,q) finds out pivot point in the range [p,q], and lifts up the current cloth to the signal level at pivot point. Then it computes new cloth using the process of Section 3.1 and calls recursively Covering() for each of left and right parts.

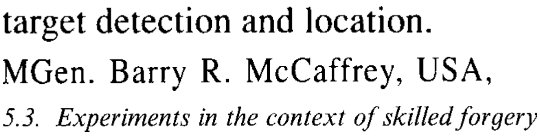
The average and worst case time complexity of Algorithm 1 is O($n$log$n$) and O(n^2) where n is the number of balls, i.e., length of signal.

Ⅳ. Application to Script Identification

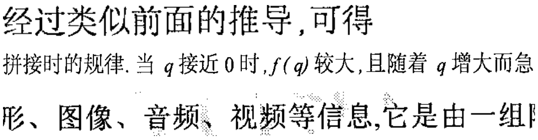
The script identification is a sub problem of OCR (Optical Character Recognition) [11]. [Figure 5] shows text line images written in different scripts, Korean, Latin, and Chinese. In conventional researches, Fourier coefficients extracted from profile and projection of line images were used as feature set [12][13]. Some literatures applied Gaussian filter to profile and projection signals before Fourier transform in order to suppress the details of signals. The aim of this section is to compare the proposed cloth covering with Gaussian filter in terms of feature's discriminating power.



(a) Korean



(b) Latin



(c) Chinese

Figure 5. Text line images written in three scripts

1. Feature extraction

One of the popularly used features is profile and projection. [Figure 6](b) presents top profile, bottom profile, and projection. The Fourier coefficients taken from Fourier transform of profile and projection

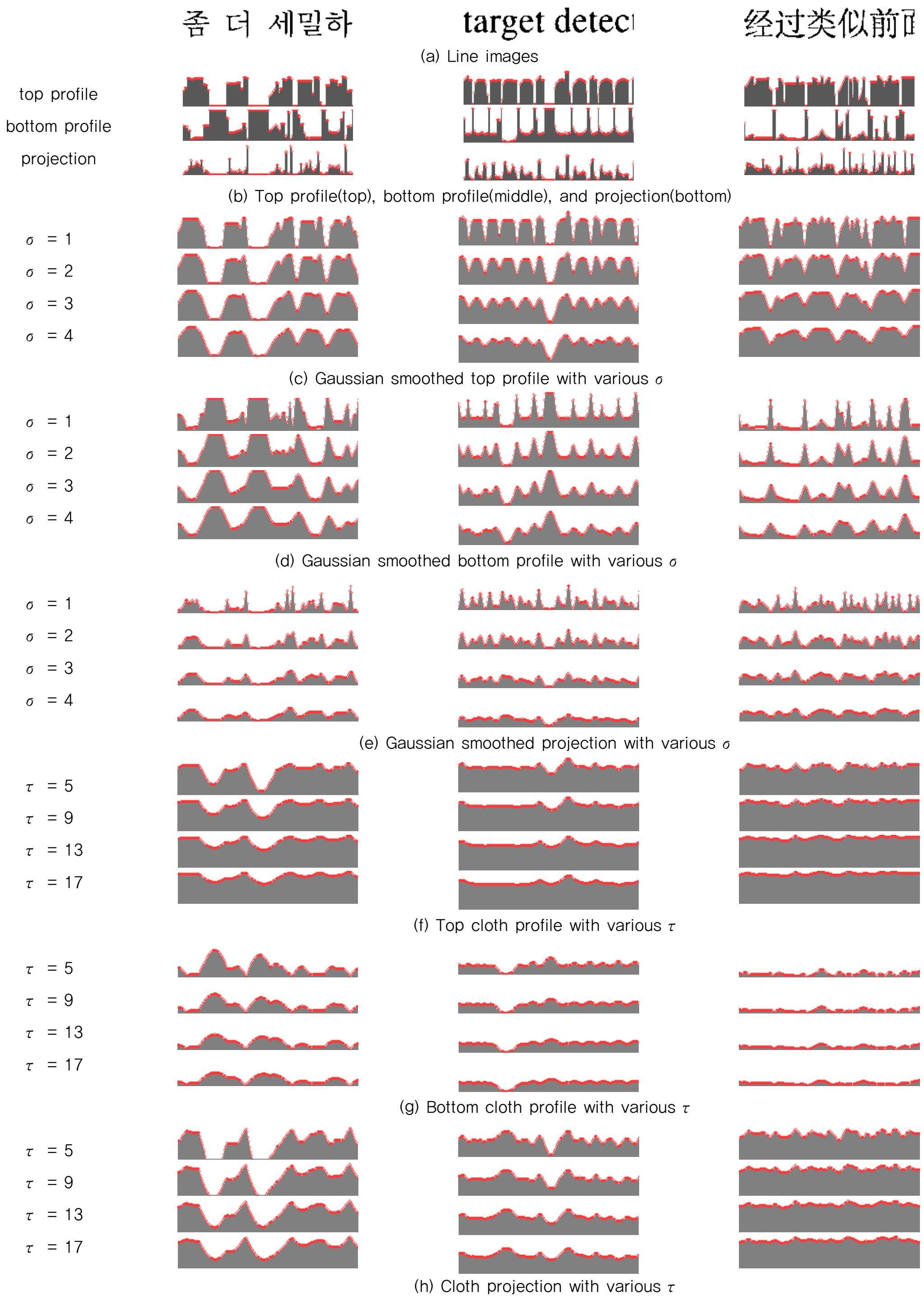


Figure 6. Top profile, bottom profile, and projection

signals were used as feature vectors by the classifier [12][13]. This paper calls this feature vector as Raw-Fourier.

As we can observe in [Figure 6](b), the raw profile and projection conveys detail shapes of input line images. We may think of using coarse version of profile and projection with the hope of improving discriminating power of the feature vector. [Figures 6](c), [Figures 6](d), and [Figures 6](e) show the Gaussian-smoothed profile and projection. Fourier coefficients taken from Fourier transform of Gaussian-smoothed profile and projection signals are used as feature vector. The feature vector is called as Gaussian-Fourier(σ) where σ represents standard deviation of Gaussian function being used.

The cloth covering can also be used to suppress the details. The result of cloth covering over the top (bottom) of line image is called as *top* (*bottom*) *cloth profile*. The difference of top and bottom profiles is called as *cloth projection*. Figures 6(f), 6(g), and 6(h) depict the cloth profile and cloth projection. Fourier coefficients taken from Fourier transform of cloth profile and cloth projection signals are used as feature vector and it is called as Covering-Fourier(τ) where τ represents stiffness coefficient.

2. Experimental results and analysis

The data sets used by experiments have been collected from various documents, including technical papers, magazines, and books. The used scripts are Korean, Latin, and Chinese. The images were scanned in 300dpi. [Table 1] shows size of data sets. For classification, we used SVM [14]. We used the polynomail kernel with the order of one (linear kernel). In measuring recognition accuracy, we used 5-fold cross validation.

Table 1. Size of data sets

	Korean	Latin	Chinese
the number of documents	69	44	37
the number of text lines	2874	2818	1337

[Table 2] presents recognition accuracy measured for three types of feature vector, Raw-Fourier, Gaussian-Fourier(σ), and Covering-Fourier(τ). To get finer analysis, we took a different number of coefficients from Fourier transform. When we took 50 coefficients from one signal, the total number of features is 150 since we used three signals, top profile, bottom profile, and projection.

Comparing raw signals and smoothed signals, the

Table 2. Recognition accuracies

Feature extraction methods		the number of features				
		150	120	90	60	30
Raw-Fourier		97.67	96.90	96.50	94.28	89.19
Gaussian-Fourier	σ = 1	97.71	97.03	96.64	94.57	89.44
	σ = 2	97.68	97.20	96.57	94.58	89.50
	σ = 3	97.81	97.37	96.63	94.58	89.60
	σ = 4	97.62	97.48	96.76	94.76	89.66
	σ = 5	97.52	97.37	96.69	94.58	89.74
Covering-Fourier	τ = 1	97.55	97.52	97.27	95.72	93.10
	τ = 5	98.41	98.26	98.01	97.52	**95.58**
	τ = 9	**98.79**	98.72	98.46	97.70	95.55
	τ = 13	98.75	98.75	**98.73**	97.72	95.08
	τ = 17	98.71	**98.78**	**98.73**	**97.79**	94.66
	τ = 21	98.71	98.72	98.62	97.52	94.05

table makes sure that both Gaussian and cloth covering are superior to raw signals. For example, when the number of features is 30, Covering-Fourier improved the accuracy of Raw-Fourier by 7.16% (from 89.19 to 95.58). We argue that using the detail shapes as they are does not provide good discriminating power of features. In each column, the best accuracy is visually emphasized using bold typeface. The cloth covering won every column. The cloth covering was better than Gaussian by more than about 1%. We recommend 9 or 13 for the scale parameter τ.

The cloth covering produced similar accuracies for the feature vector size of 90~150. On the contrary, Raw-Fourier and Gaussian-Fourier decreases rapidly as the number of features decreases.

3. Discussions

In order to improve the reliability of experimental results, we conducted a statistical significance test. [Table 3] shows the t-test result.

Table 3. Result of statistical significance test

	the number of features				
	150	120	90	60	30
F-test	0.006	0.03	0.0006	0.0003	0.0006
t-test	0.006	0.001	0.0009	0.0004	0.00005

We can ensure that the proposed Covering-Fourier is superior if the null hypothesis is rejected by the two-tailed test with significance level α=0.05 under the hypothesis H_0:$\mu_1-\mu_2=0$, H_a:$\mu_1-\mu_2>0$. For the t-test, we first performed F-test to identify whether two populations (Gaussian-Fourier and Covering-Fourier) have equal variance or unequal variance. As we can see in [Table 3], for all of the columns, F-test showed that the results had smaller values than the significance level α=0.05. So the unequal variance t-test was performed. The t-test resulted in smaller values than α=0.05 for all the columns. So the null hypothesis was rejected. We concluded that the Covering-Fourier is superior than the Gaussian-Fourier.

We also attempted to explain the reason of superiority analytically. [Figure 7] illustrates a simple situation of input signal which could be used to explain the property of distinctiveness of the features used in the experiments. The Gaussian mask with size 3, (0.25,0.5,0.25) will produce a flat signal which has the value 1.5h all over the whole positions. The result is the same as the input signal which has a constant value, 1.5h. On the contrary, the cloth covering might produce a result signal which preserves the rising-and-falling patterns of the original signal.

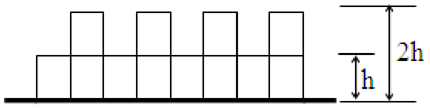


Figure 7. A situation illustrating distinctiveness of features

The [Figure 7] is just a simple situation for the purpose of showing the difference of cloth covering and Gaussian. However we can make other situations which produce similar results for larger masks such as with size 5, i.e, (0.1,0.2,0.4,0.2,0.1). We may conclude that the cloth covering is a kind of *structure-preserving* operation while Gaussian is not. We believe that this property is a source of better discrinative power of the cloth covering.

V. Conclusions

This paper proposed a novel multi-scale method called cloth covering. The method has lots of potentioal application areas. The paper presented algorithms for 1-D cloth covering. Using script recognition experiments, we showed viability of the cloth covering. For an objective comparison, we performed t-test. Also we described the reason of the method being sueprior to the Gaussian kernel. Until now, the performance test has been done only for the script identification problem. Expanding to other applications is one of major future works.

There are a number of futures of the cloth covering. Firstly, mathematical properties are worth of studying. Comparison with Gaussian and mathematical morphology could guide the study. Secondly, application areas in which the cloth covering competes with or superior to conventional method should be identified. As a specific area, keypoint detection for SIFT is being studied by authors. The conventional SIFT uses DOG (Difference of Gaussian). Our concern is to test the performance when we replace DOG with DOC (Difference of Cloth covering). Thirdly, algorithms for n-dimensional signals should be developed. The priority is on developing algorithm for 2-D images.

참 고 문 헌

[1] T. Lindeberg, "Scale-space theory: a basic tool for analyzing structures at different scales," Journal of Applied Statistics, Vol.21, No.2, pp.224-270, 1994.

[2] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 3rd Ed., Pearson, 2010.

[3] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, Vol.60, No.2, pp.91-110, 2004.

[4] A. L. Yuille and T. A. Poggio, "Scaling theorems for zero crossings," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.8, pp.15-25, 1986.

[5] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection," International Journal of Computer Vision, Vol.30, No.2, pp.117-156, 1998.

[6] K. Mikolajczyk and C. Schmid, "Scale and affine invariant interest point detectors," International Journal of Computer Vision, Vol.60, No.1, pp.63-86, 2004.

[7] S. G. Mallat, *A Wavelet Tour to Signal Processing*, Academic Press, 1999.

[8] D. Luebke, *Level of Detail for 3D Graphics*, Elsevier science, 2003.

[9] Y. Lee, S. Yoon, S. Oh, D. Kim, and S. Choi, "Multi-resolution cloth simulation," Pacific Graphics, Vol.29, No.7, 2010.

[10] D. Halliday, *Fundamentals of Physics*, 8th Ed. Wiley, 2008.

[11] S. Marinai, *Machine Learning in Document Analysis and Recognition*, Springer, 2008.

[12] D. Ghosh, T. Dude, and A. P. Shivaprasad, "Script recognition - a review," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.32, No.12, pp.2142-2161, 2010.

[13] U. Pal and B. B. Chaudhuri, "Identification of different script lines from multi-script documents," Image and Vision Computing, Vol.20, pp.945-954, 2002.

[14] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," Data Mining and Knowledge Discovery, Vol.2, pp.121-167, 1998.

저 자 소 개

김 민 우(Minwoo Kim) 정회원

- 2007년 8월 : 전북대학교 전자정보공학부(컴퓨터)(공학사)
- 2009년 8월 : 전북대학교 전자정보공학부(컴퓨터)(공학석사)
- 2009년 9월 ~ 현재 : 전북대학교 전자정보공학부(컴퓨터) 박사과정

<관심분야> : 컴퓨터비전, 패턴인식

오 일 석(Il-Seok Oh) 정회원

- 1984년 : 서울대학교 컴퓨터공학과(공학사)
- 1992년 : KAIST 전산학과 박사
- 1992년 ~ 현재 : 전북대학교 컴퓨터공학부 교수

<관심분야> : 컴퓨터비전, 패턴인식