

# Method for Extracting Product Information from TV Commercial

Kohei Arai

Information Science Department  
Saga University  
Saga, Japan

Herman Tolle

Software Engineering Department  
Brawijaya University  
Malang, Indonesia

**Abstract** — Television (TV) Commercial program contains important product information that displayed only in seconds. People who need that information has no insufficient time for noted it, even just for reading that information. This research work focus on automatically detect text and extract important information from a TV commercial to provide information in real time and for video indexing. We propose method for product information extraction from TV commercial using knowledge based system with pattern matching rule based method. Implementation and experiments on 50 commercial screenshot images achieved a high accuracy result on text extraction and information recognition.

**Keywords** - text detection; information extraction; rule based classifying; patern matching;

## I. INTRODUCTION

Nowadays, people use the information from Television (TV) commercial program as a reference before they buy the product. Since TV commercial only displays the important information in seconds, people who need such information has no insufficient time for noted it or even just for reading that information. In Japan, there is a company provide services for reading, note and distribute such information from TV commercials, but they still did it in manually. If these text occurrences could be detected, segmented and recognized automatically, it would be a valuable source for information extraction, indexing or retrieval.

The text information extraction (TIE) problem can be divided into the following sub-problems: (i) detection, (ii) localization, (iii) tracking, (iv) extraction and enhancement, and (v) recognition [1]. A TIE system receives an input in the form of a still image or a sequence of images. The images can be in gray scale or color, compressed or un-compressed, and the text in the images may or may not move. Text detection refers to the determination of the presence of text in a given frame (normally text detection is used for a sequence of images). Text localization is the process of determining the location of text in the image and generating bounding boxes around the text. Text tracking is performed to reduce the processing time for text localization and to maintain the integrity of position across adjacent frames. Although the precise location of text in an image can be indicated by bounding boxes, the text still needs to be segmented from the background to facilitate its recognition. This means that the extracted text image has to be converted to a binary image and enhanced before it is fed into

an optical character recognition (OCR) engine. Text extraction is the stage where the text components are segmented from the background. Enhancement of the extracted text components is required because the text region usually has low-resolution and is prone to noise. Thereafter, the extracted text images can be transformed into plain text using OCR technology. Text in videos is usually not easily extracted especially when it is embedded in complex background scenes and suffers from poor visual quality due to the effects of motion blur and compression artifacts.

After extracting and recognition text from TV commercial screenshot, we should identify and classify what type of information from extracted text. Important product information from TV commercials is product name, product price, URL information, and phone number. In this paper we propose a novel method for extracting product information from TV commercials by using knowledge based method with pattern matching and classification rules for recognizing and classifying the information.

## II. RELATED WORK

In text extraction from image, text localization methods can be categorized into three types: region-based, texture-based and hybrid approaches. Region-based schemes use the properties of color or grayscale in a text region or their differences with corresponding properties of background. These methods can be divided further into two sub-approaches: connected component (CC) and edge-based [2].

CC-based methods apply a bottom-up approach by grouping small components into successively larger ones until all regions are identified in the image. A geometrical analysis is required to merge the text components using the spatial arrangement of the components so as to filter out non text components and mark the boundaries of text regions. Among the several textual properties in an image, edge-based methods focus on the 'high contrast between the text and the background'. The edges of the text boundary are identified and merged, and then several heuristics are performed to filter out the non text regions.

Texture-based methods use the observation that text in images has distinct textural properties that distinguish them from the background. The techniques based on Gabor filters, Wavelet, FFT, spatial variance, etc. can be performed to detect the textural properties of a text region in an image [1]. First

class of localization methods can be found in some works [3] – [6]. Gllavata et al. [3] have presented a method to localize and extract text automatically from color images. First, they transform the color image into grayscale image and then only the Y component is used. The text candidates are found by analyzing the projection profile of the edge image. Finally, a *binarized* text image is generated using a simple binarization algorithm based on a global threshold. They [4] also have applied the same idea of the previously mentioned paper to localize text; in addition the algorithm has been extended with a local *thresholding* technique.

Cai et al. [5] have presented a text detection approach which uses character features like edge strength, edge density and horizontal alignment. First, they apply a color edge detection algorithm in YUV color space and filter out non text edges using a low threshold. Then, a local *thresholding* technique is employed to keep low-contrast text and further simplify the background. An image enhancement process using two different convolution kernels follows. Finally, projection profiles are analyzed to localize the text regions. Jain and Yu [6] first employ color reduction by bit dropping and color clustering quantization, and afterwards a multi-value image decomposition algorithm is applied to decompose the input image into multiple foreground and background images. Then, CC analysis is performed on each of them to localize text candidates. From all method described here, there is no method specific for TV commercial except Leinhart [9] and Gllavata [3].

### III. THE PROPOSED METHOD FOR TEXT EXTRACTION FROM TV COMMERCIALS

#### A. Overview of Information Extraction from TV Commercial System

Information extraction from TV commercial is useful for help people recognize important information from short and fast commercials video. The block diagram of TV commercial Information extraction system is shown in Figure 1 consist five processes as follow: Video frame detection, Text detection, Text Localization, Text Extraction, Text Recognition and Text Identification. Text detection in TV commercial video should perform based on the characteristic of displayed text object in commercials. After the investigation on common TV commercials video, we found some typical pattern of the commercials that accepted as assumption in our research work, as follow:

- 1) Text information usually in bright or contrast color like: white, red, or yellow in dark background; or blue or black in bright background
- 2) Important information using bigger font size.
- 3) Important information usually appears in near end part of commercials.

For our text extraction process, we propose a new text extraction method using combination of edge-based method and CC-based method. We use edge-based method for text detection because the characteristic of commercial video which is contain contrast color of text with background. We use CC-based method for text localization because we need the information about text position and size that important in

information recognition. Also both of two method has less complexity in algorithm for real time application,

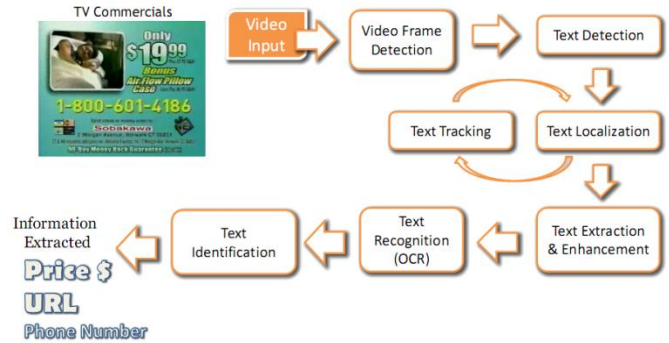


Figure 1. Block Diagram of Automatic TV Commercial Information Extraction

#### B. Video Frame Detection

By investigation on common TV commercial video samples, we found some typical pattern of the commercials that accepted as assumption in our research work, as follows:

1. Video length in 15 seconds, 30 seconds, 45 seconds or 60 seconds
2. Text information usually in bright or contrast color: Some pattern is: white, orange, or yellow in dark background, in the other hand, black, red or blue in bright background.
3. Important information is displayed more than once among a TVCM.
4. Important information using bigger font size.
5. Important information appears in the near of the end part of a TVCM.
6. Some important information displayed with product image.

Information extraction from TV commercial video is based on text extraction from screenshot image of commercials video. Commercial video was recorded from television and classify based on the video length. Then in determined time, we capture the screen shoot of running CM In automatic way, we can implement other research work on commercial detection like [14] for detecting the occurrence of commercials within a TV programs. In this paper, we assume that commercials are already separated from other TV programs. Using assumption that important information usually appears in near end part of commercials we generate screenshot of commercial video repeated in every 3 second but vary depend on commercials long. We use 15 seconds and 30 seconds type for TVCM. The screenshot number and its relationship with video frame are shown in Table 1.

TABLE I. TIMING FOR SCREENSHOT TAKING AND NUMBER OF SCREENSHOT

CM Duration	Seconds of Screenshot	Total
15 seconds	5, 8, 11, 14	4
30 seconds	14, 17, 20, 23, 26, 29	6

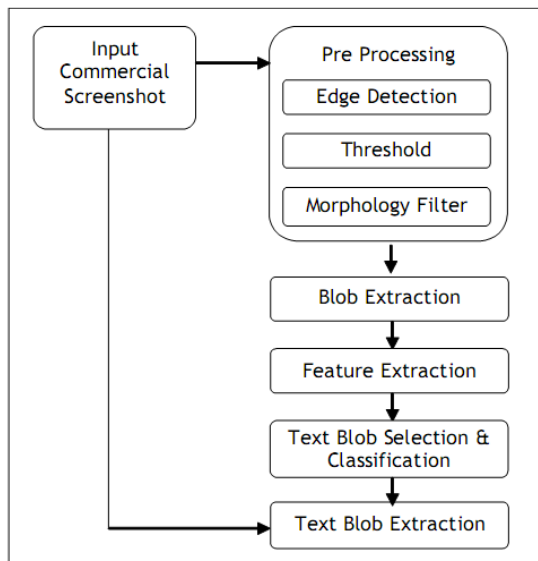


Figure 2. Text Extraction Algorithm

### C. Text Detection Algorithm

The text detection stage seeks to detect the presence of text in a given image. The text detection method comprises of four main phases; pre processing, blob extraction, blob classification, and text blob extraction. The block schematic of the text extraction methodology is given in Figure 2. The detailed description of each phase is presented in the following subsections. This method also published in [12].

#### B. Pre-Processing

The objective of pre-processing is to convert image as binary picture that separates text object from background or non text object. We use combination of filter to process original image into binary image. Processes within pre-processing process as follows:

1. Extract Red channel.
2. Convert image into Grayscale
3. Homogeneity Edge Detection Filter
4. Binarization with threshold
5. Morphology Erosion Filter
6. Dilatation Filter

First, we extract red channel from image to get bright image information. After image is converted to grayscale and extract red channel, we implement homogeneity edge detection filter to get edge pattern of the object from image. The next process is binarization –after inverting image color- with an appropriate threshold number to produce black and white image. The heuristic value of threshold is 225 (for the images with quantization bits of 8 bits) that chosen empirically based on experiments. Combination of edge detection filter and appropriate threshold number will separates text from relatively complex background. In the end of pre-processing, we implement morphology erosion filter in 5x5 horizontal matrixes to combine small blob with left or right nearest blob. Figure 3 show step-by-step results from process in pre-processing process.

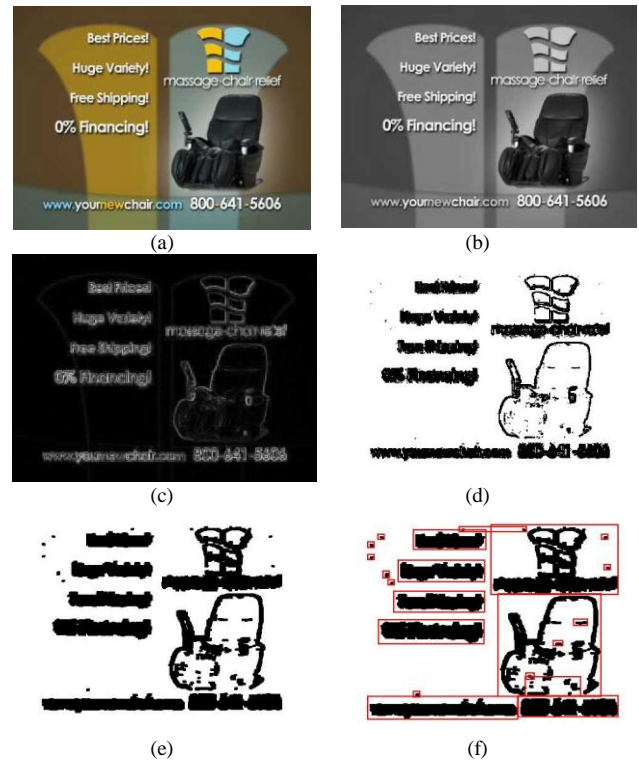


Figure 3. Process in pre-processing, Original (a), Grayscale (b) Homogeneity Edge detection (c), Binarization (d), Morphology Erosion (e), Blob Extraction (f).

#### C. Blob Extraction

After pre-processing, we detect all connected pixels with same color index as a separate blob object. Blob extraction is done using same blob extraction function implemented in our Comic text extraction method [11]. This process produces text blobs and also non text blobs. To classify a blob as a text blob or non text blob, we extract some features from text blob candidate for classification. In text blob extraction using blob extraction function, we select only blob with minimal size that selected as candidate of text blob. The minimal size of the text candidate blob width is  $[Image.Width]/40$  and the text blob height is  $[Image.Height]/40$ . Parameter of  $[Image]$  corresponds to input image size. Figure 3.f show sample of detected blob in a commercial image after pre-processing that contain text blob, non-text blob and noise blob. Then we implement text blob selection to select only text blob and to remove noise or non-text blob, and text blob classification to classify blob into text word or sentence based on text blob position.

$$TB[i].YCenter = TB[i].Top + (TB[i].Height/2) \quad (1)$$

Where;

$TB[i]$  corresponds to  $i^{th}$  blob of detected text blob.

$TB[i].Left$  corresponds to left point position parameter of detected text blob

$TB[i].Width$  corresponds to width parameter of detected text blob

For selecting and classifying text blob into horizontal text region, we use some rules for classification as follows:

1. Blob size is smaller than half of image size.
2. Classify all blobs with same vertical center point and relatively same height size into the same cluster. Assuming difference distance between centers is  $BD_{min}$  and difference size between 2 blobs is  $BS_{min}$ .

$$\text{If } (TB[i].XCenter \pm BD_{min}) \in \text{Column}[j] \text{ Then } TB[i] \text{ member of Column}[j] \quad (2)$$

3. Position of blob center Y ( $Blob.Cy$ ) is in range ( $Blob[i].Cy - Blob[i].Height/2 < Blob[j].Cy$ ) and ( $Blob[i].Cy + Blob[i].Height/2 > Blob[j].Cy$ ) (3)
4. Ignore blob cluster with width size smaller than its height size.

Minimal distance  $BD_{min}$  is approximately less than half of average text blob height and  $BS_{min}$  is around 40% of difference width between two blobs. Figure 4 Show the sample of text detection process with text blob detection (4.b) and text blob classification (4.c)

#### D. Text Extraction and Enhancement

After all text blob candidate localized, then extracted into separate text blob candidate. Extraction blob is getting from original input image without any other processing. We extract every blob from original image using position information (top, left, width, height) of the blob.



Figure 4. Sample of Commercial Screenshot (a), Blob detection (b), Text blob classification (c),

After all blob extracted, then we should enhance the text blob before using OCR for recognizing the text. We implement simple text enhancement with binarization (threshold) using Otsu [13] method. Otsu method tries to find minima between two peaks in histograms. Otsu's method we exhaustively search for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes. Figure 5 shows the sample of original extracted text blob and after pre-processing for enhancement.



Figure 5. Sample of text blob originally extracted (left), and with pre-processing (right)

#### IV. KNOWLEDGE BASED METHOD FOR INFORMATION RECOGNITION

For recognizing and classifying the information from extracted text, we design a TV commercial information extraction knowledge based system with rules and pattern matching. We design a knowledge based system with specific pattern for each type of information that we want to extract. The important information to extract from TVCM is: phone number, URL information, price information and product name. After detected and extracted text from TVCM screenshot, we should extract the ASCII text representation of image text using OCR application. The accuracy of OCR results is depending on OCR application that not covers in this paper. Then by using knowledge based system we try to classify the information based from rules in knowledge based.

We should extract some features from text blob before put in selection process by using knowledge based system. The features should be extracted for each blob word as follows: text ASCII representation, text blob position (top, left), text blob size (width, height), and relative position (top | middle | bottom and left | center | right), character type (text, number, currency, or URL keyword). Then, by using data of extracted features we can classify text based on our knowledge based rules and pattern.

Figure 6 shows some example of extracted text blob from screenshot of TV commercial video.



Figure 6. Sample of extracted text blob from text extraction process

The knowledge based rule system and specific pattern for each type of information is described in the following sections:

A. Rules for Phone Number:

Definition rules for selection a text combination as a phone number is:

- Surround with text related to phone information: Call | Now | Phone
- Should only contain: numbers [0-9] and optional for: hyphen [-]

For Phone Type 1:

- Match the string pattern: xxxx-xxx-xxx where x is numbers.
- Match the first 4 number combinations: 0120 (Japan).
- Total numeric string number is 10 without hyphen.
- Minimal Two hyphen (-) detected.

For Phone Type 2:

- Match the string pattern: x-xxx-xxx-xxxx where x is numbers
- Match the first 4 number combinations: 1-800 (Japan)
- Total numeric string number is 11 without hyphen.
- Three hyphen detected

B. Rules for Price Information:

Definition rules for selection a text combination as price information is:

- Match the string pattern: x,xxx or xx,xxx or xxx,xxx
- Should only contain numbers [0-9], and optional for [.] [,]
- String size is relatively large
- Surround with text related to price information: [Only | Price | Now ]
- Surround with character related to currency sign [\$, \]

C. Rules for Web Site Information:

Definition rules for selection a text combination as web site information is:

- Match the string pattern (only one, combination or all): [http:// | www. | .com | .net | .org | .jp ]
- Web site URL name check function

D. Rules for Product Name:

Definition rules for selection a text combination as web site information is:

- String size is relatively larger than others
- Position is relatively in center top
- Found more than x times in video frame screenshot, while x > 50% of processing frame number in one commercial.

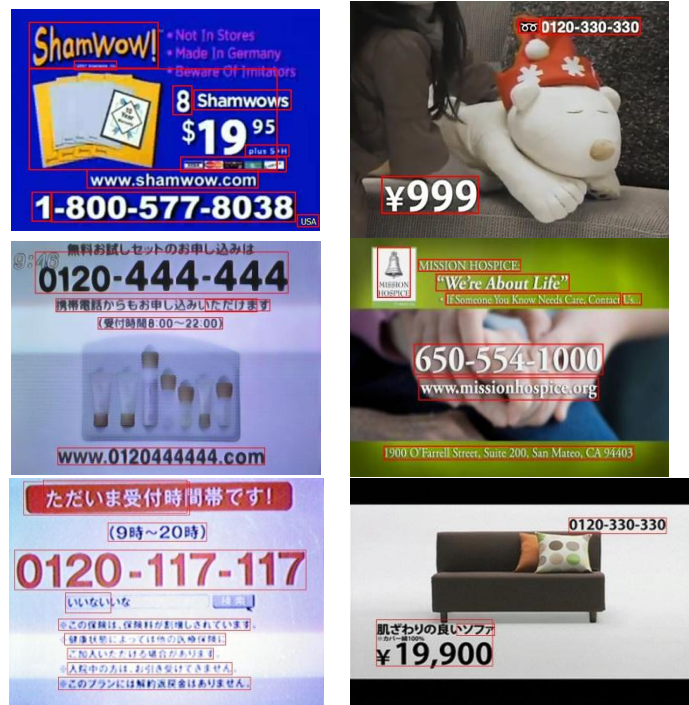


Figure 7. Sample of TV Commercial Screenshot Image with detected text in red rectangle

V. EXPERIMENTAL RESULTS

We implement our method for text detection and extraction based on AForge open source image processing tools running on Visual Studio 2008. The proposed approach has been evaluated through data set containing different type of commercial screenshot images taken from Japan TV and some TV commercial screenshot randomly from internet. The based blob extraction method is the same blob extraction function we use in our text extraction method from digital Manga comic [11]. The performance of the system is evaluated on the text box level in terms of precision and recall. We also develop system for information recognition and evaluating it in real results from OCR and in ideal condition assuming OCR has no error on recognition the text. Figure 7 show the samples of TV commercial screenshot images with detected text located in red rectangle.

A. Evaluation of Text Extraction

The results for the experiments on text extraction are summarized in Table 2 where the number of existing text lines, the number of detected text lines, the number of false alarms and the corresponding values for recall and precision are listed. We use about 50 screenshot images from different TV commercial scene from image with simple text on background to text with complex background.

TABLE II. EXPERIMENTAL RESULT FOR TEXT EXTRACTION.

Number of images	50
Number of text lines	250
#correct detected	243
#false positives	19
Recall (%)	92.75%
Precision (%)	97.20%

Recall is defined as:

$$Precision = \frac{Correct\ Detected}{(Correct\ Detected + Missed\ Text\ Lines)}$$

Whereas precision is defined as:

$$Recall = \frac{Correct\ Detected}{(Correct\ Detected + False\ Positives)}$$

A text line is considered as detected correctly, if a text line consists in an extracted blob text, while a detected text blob is considered as a false alarm, if no text appears in that extracted text. The text extraction algorithm achieved a recall of 92.75% and a precision of 97.20%. The precision of our method is relatively better than Leinhart's [9] method and Gllavata's [3] method for detection text on commercial video as shown in Table 3.

TABLE III. COMPARISON OF PRECISION AND RECALL OF OUR METHOD WITH OTHER METHODS.

Method	Recall (%)	Precision (%)
Lienhart and al.Method [9]	-	66.00%
Gllvata and al.Method [3]	87.00%	90.00%
Our Method	92.75%	97.20%

### B. Evaluation of OCR & Text Enhancement

Implementation of the character recognition process currently still uses commercial OCR application. We use Softi free OCR application. We evaluate the accuracy of character recognition without enhancement and with enhancement. We recognize only roman character within commercial text candidate blobs, and ignored Japanese character recognition at this time. There are around 750 characters in 50 samples of images, excluding space character. The results in Table 4 shown that by implementing the enhancement process, we can improve about 10% accuracy of the recognition.

TABLE IV. EVALUATION RESULTS OF CHARACTER RECOGNITION WITH/WITHOUT ENHANCEMENT

	Without Enhancement	With Enhancement
Total Character	748	748
Correct Detection	565	636
Miss Detection	183	112
Accuracy (%)	<b>75.53</b>	<b>85.03</b>

### C. Evaluation of Information Extraction Knowledge Based Systems

Evaluation of the accuracy of knowledge based rule system is conduct through 2 conditions. First is real condition with real results from OCR, and second condition is ideal condition while assuming that we have 100% accuracy results from OCR. While first condition evaluate for the whole systems, second condition is only for evaluate the accuracy of knowledge based rule system assuming that there is no error from OCR. Table 5 show samples of text blob, its extracted information and type of information by the knowledge based rules. Table 6 shows the results of the accuracy evaluation of knowledge based. The term  $Tot_1$  means the total of product

information data from real OCR; the term  $Tot_2$  means the total of product information data from perfect OCR; the term C means the correct detection; the term M means the miss detection; and the term FP means the False Positive. The accuracy of knowledge based system (only) for: *phone number* recognition: 86.36%; *URL address* recognition: 78.57% and *price information* recognition: 73.33%. And the accuracy of knowledge based system in whole system with original condition of OCR is: 75% accuracy of *phone number* recognition; 70% accuracy of *URL address* recognition and 60% accuracy of *price information*.

TABLE V. SAMPLE OF TEXT BLOB INFORMATION CLASSIFICATION.

Text Blob	Extracted Information	Type of Inf.
0120-330-330	0120-330-330	Phone
www.drylandcorp.com	www.drylandcorp.com	URL
¥1,900	¥1.900	Price
www.0120444444.com	www.0120444444.com	URL
Self Medication	Self Medication	-
954-534-4573	954-543-4573	Phone
nissen.jp	nissen.jp	URL
mightywallets.com	mightywallets.com	URL
9:48	9:48	-
1-8 September	1-8 September	-
\$15.00 with Free Shipping*	\$15.00 with Free Shipping	Price

TABLE VI. EVALUATION RESULTS OF KNOWLEDGE BASED RULE SYSTEM

Information	$Tot_1$	C	M	FP	$Tot_2$	C	M	FP
Phone Number	16	12	4	2	22	19	3	0
URL Address	10	7	3	1	14	11	3	0
Price Information	10	6	4	2	15	11	4	1

### D. Discussion

Based on our experimental results, accuracy of recall on text extraction is not so high because the occurrence of many false positive, that is some non text objects are detected as text in our approach. Since the text blob candidate will then send to OCR process for recognition, non-text object is not a significant problem, because non-text object has no text output from OCR process. Also, if OCR generated results for non-text object, usually only a little text with no meaning that can be ignored. Although our experimental results only calculate the text from commercial video screenshot image, it is still possible to implementing in real time analysis for TV commercial video.

The accuracy results of information recognition are also depending on the accuracy of OCR application. From the evaluation process while assuming using the perfect OCR, knowledge based system accuracy for product phone number is about 86.36%, it is mean that our method for information recognition knowledge based rules and pattern matching method should be improved for better results.

## VI. CONCLUSION & FUTURE WORK

In this paper, we have proposed an approach to automatically extracted text appearing in commercial screenshot images and recognize the product information based on the pattern matching method. We see that our method on

text extraction have good performance for localization of text in commercial image, also our information recognition has high accuracy on classifying and recognizing product information from TV commercials. This is notable that in the future, we can improve the method for implementing a real time product information extraction from TV commercial application, for using in a set top box TV system, as an application for helping people automatically retrieve important information from live streaming TV video contents.

#### REFERENCES

- [1] K. Jung, K.I. Kim, and A.K. Jain, "Text Information Extraction in Images and Video: A Survey," *Pattern Recognition*, Vol. 37, pp. 977-997, 2004.
- [2] G. Aghajari, J. Shanbehzadeh, and A. Sarrafzadeh, "A Text Localization Algorithm in Color Image via New Projection Profile", *Proceedings of International MultiConference of Engineers and Computer Scientists 2010*, Vol II, pp. 1486-1489, 2010
- [3] J. Gllavata, R. Ewerth, and B. Freisleben, "A Robust Algorithm for Text Detection in Images", *Proc. of 3rd Int'l Symposium on Image and Signal Processing and Analysis*, Rome, pp. 611-616, 2003.
- [4] J. Gllavata, R. Ewerth and B. Freisleben, "Finding Text in Images via Local Thresholding", *International Symposium on Signal Processing and Information Technology*, Darmstadt, Germany, pp. 539-542, 2003.
- [5] M. Cai, J. Song, and M. R. Lyu, "A New Approach for Video Text Detection", *Proc. of IEEE Int'l Conference on Image Processing*, Rochester, New York, USA, pp. 117-120, 2002.
- [6] A. K. Jain, and B. Yu, "Automatic Text Location in Images and Video Frames", *Pattern Recognition* Vol 31, No. 12, pp. 2055-2076, 1998.
- [7] Q. Ye, Q. Huang, W. Gao and D. Zhao, "Fast and robust text detection in images and video frames", *Image Vision Comput.* 23, pp. 565-576, 2005.
- [8] J. Gllavata, R.Ewerth and B. freisleben, "A text detection, localization and segmentation system for OCR in images," in *Proc. IEEE Sixth International Symposium on Multimedia Software Engineering*, pp. 425-428, 2004.
- [9] Rainer Lienhart and Wolfgang Effelsberg, "Automatic text segmentation and text recognition for video indexing". *Multimedia Syst.* Vol 8, No. 1, pp 69-81, 2000.
- [10] Xiaojun Li, Weiqiang Wang, Shuqiang Jiang, Qingming Huang, Wen Gao, "Fast and Effective Text Detection", *IEEE International Conference on Image Processing*, San Diego, USA, pp. 969-972, 2008.
- [11] Kohei, A., Tolle, H., "Method for Real Time Text Extraction from Digital Manga Comic", *International Journal of Image Processing* Vol 4, No. 6, pp. 669-676, 2011.
- [12] Kohei Arai, Tolle Herman, "Text Extraction from TV Commercial using Blob Extraction Method", *International Journal of Research Review on Computer Science (IJRRCS)*, Vol 2, No.3, pp. 895-899, 2011.
- [13] N. Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics.* 1979, 9(1): 62-66
- [14] Yijun Li, Danqing Zhang, Xiangmin Zhou, and Jesse S. Jin. A confidence based recognition system for TV commercial extraction. In *Proceedings of the nineteenth conference on Australasian database - Volume 75 (ADC '08)*, Vol. 75. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 57-64. 2008.
- [15] Pritam Singh Negi, M M S Rauthan and H S Dhami. Article: "Text Summarization for Information Retrieval using Pattern Recognition Techniques". *International Journal of Computer Applications* 21(10):20-24, 2011.
- [16] Ralph Grishman, "Information Extraction: Techniques and Challenges", . In *Information Extraction: A Multidisciplinary Approach to an Emerging Information Technology*, in M.T. Pazzienza, (ed.), Springer, 97.

#### AUTHORS PROFILE



Kohei Arai, He received BS, MS and PhD degrees in 1972, 1974 and 1982, respectively. He was with The Institute for Industrial Science and Technology of the University of Tokyo from April 1974 to December 1978 and also was with National Space Development Agency of Japan from January, 1979 to March, 1990. During from 1985 to 1987, he was with Canada Centre for Remote Sensing as a Post Doctoral Fellow of National Science and Engineering Research Council of Canada. He moved to Saga University as a Professor in Department of Information Science on April 1990. He was a councilor for the Aeronautics and Space related to the Technology Committee of the Ministry of Science and Technology during from 1998 to 2000. He was a councilor of Saga University for 2002 and 2003. He also was an executive councilor for the Remote Sensing Society of Japan for 2003 to 2005. He is an Adjunct Professor of University of Arizona, USA since 1998. He also is Vice Chairman of the Commission A of ICSU/COSPAR since 2008. He wrote 26 books and published 227 journal papers.



Herman Tolle, He graduated Bachelor degree in Electrical Engineering from Brawijaya University, Malang in 1998, and also graduated Master degree in Telecommunication Information System from Bandung Institute of Technology (ITB), Bandung in 2002. He is with Engineering Faculty of Brawijaya University from 2002 to present. He is now a Doctoral student in Department of Information Science, Faculty of Science and Engineering, Saga University Japan. He has major concern of research in image analysis, multimedia, content adaptation and web engineering.