A Content-based Image Retrieval of Muay-Thai Folklores by Salient Region Matching

Pakpoom Mookdarsanit[1], Lawankorn Mookdarsanit[2]

[1]Department of Computer Science
Faculty of Science, Chandrakasem Rajabhat University
Bangkok, Thailand
pakpoom.m@chandra.ac.th
[2]Department of Business Computer
Faculty of Management Science, Chandrakasem Rajabhat University
Bangkok, Thailand
lawankorn.s@chandra.ac.th

บทคัดย่อ—บทความนี้นำเสนอการค้นคืน “ท่ารำมวยไทย (หรือแม่ไม้มวยไทย)” โดยใช้เทคนิคการเปรียบเทียบความเหมือนกันระหว่างเส้นขอบบริเวณของวัตถุ ซึ่งเป็นงานวิจัยฉบับแรกที่มีการบูรณาการระหว่างศาสตร์ “การจำแนกท่าทางของมนุษย์” กับ “แม่ไม้มวยไทย” โดยมีวัตถุประสงค์เพื่ออนุรักษ์ศิลปะและวัฒนธรรมไทย อย่างที่เราทราบกันดีว่า มวยไทยนั้นได้รับขนานนามว่าเป็นศิลปะการต่อสู้ที่ใช้ในการต่อสู้ได้จริงมากที่สุดในโลก ดังจะเห็นได้ว่ามีชาวต่างชาติจำนวนมากมาใช้ชีวิตในช่วงวันหยุดยาวเพื่อร่ำเรียนวิชามวยไทย จากจารึกจดหมายเหตุแห่งสยาม พบว่า แม่ไม้มวยไทยนั้นเริ่มมีการเผยแพร่ให้กับกองทัพและราษฎรของสมเด็จพระสุริเยนทราธิบดีแห่งอโยธยา เพื่อเป็นวิชาการต่อสู้ป้องกันตัวเองจากการโจมตีของศัตรูภายนอกราชอาณาจักร ซึ่งประกอบไปด้วย แม่ไม้มวยไทย 15 ท่ารำ ดังนี้ (1.) ท่าสลับฟันปลา (2.) ท่าปักษาแหวกรัง (3.) ท่านชวาซัดหอก (4.) ท่าอิเหนาแทงกฤช (5.) ท่ายกเขาพระสุเมรุ (6.) ท่าตาเถรค้ำฝัก (7.) ท่ามอญยันหลัก (8.) ท่าปักลูกทอย (9.) ท่าจระเข้ฟาดหาง (10.) ท่าหักงวงไอยรา (11.) ท่านาคาบิดหาง (12.) ท่าวิรุฬหกกลับ (13.) ท่าดับชวาลา (14.) ท่าขุนยักษ์จับลิง และ (15.) ท่าหักคอเอราวัณ ตามลำดับ

คำสำคัญ: การรู้จำท่ามวยไทย ,การค้นคืนศิลปะการป้องกันตัว, คำอธิบายภาพกีฬาการต่อสู้, แม่ไม้มวยไทย, การจำแนกท่าทางของมนุษย์

Abstract—This paper introduces a novel “Muay-Thai folklore image retrieval” by salient region matching that is the first groundwork to combine the “human action classification” and “Muay-Thai folklores” together. The paper is proposed to support the “Thai arts and cultures conservation (known as Thai-ness)” As Muay-Thai is denominated as the “real-world practice of martial fighting” in the world. Importantly, many foreigners take Muay-Thai training courses during their long holidays. From the Siamese archives, the “original Muay-Thai folklores” was disseminated in the reign of king Suriyentharathibodhi during the age of Ayutthaya. Since the king needed his men and armies having the martial fighting skills to protect the domestic kingdom against the foreign enemies. And the traditional Muay-Thai folklores consisted of (1.) Sa Lub Fun Pla, (2.) Pak Sa Waek Rang, (3.) Java Sud Hok, (4.) Tang Krich, (5.) Yok Khao Phra Sumen, (6.) Ta Then Kum Fak, (7.) Morn Yan Lak, (8.) Pak Look Toy, (9.) Jor Ra Kay Fad Harng, (10.) Hak Nguang Ayara, (11.) Na Ka Bid Harng, (12.) Wi Roon Hok Klub, (13.) Dub Chavala, (14.) Koon Yak Jub Ling and (15.) Huk Khor Erawan, respectively.

Keywords-Muay-Thai Recognition; Martial Arts Retrieval; Fighting Sports Annotation; Muay-Thai Folklores; Human Action Classification

I. INTRODUCTION

“Muay-Thai (or Thai-boxing)” is one of the most popular martial arts in the world [1], like other fighting sports as “Kung Fu”, “Judo” and “Taekwondo”. With this in mind, the uniqueness of Muay-Thai is using the physical stand-up striking together with various clinching techniques that combines the use of 2-fists, 2-elbows, 2-knees, and 2-shins [2] to fight efficiently against the opponent during the combat boxing, being known as “the art of eight limbs (Thai: นวอาวุธ)”. From the ancient Siamese archives, Muay-Thai is referred to the king Suriyenthrathibodi of Ayutthaya

kingdom [3]. Since the king helpfully published the Muay-Thai scriptures to his men and armies in order to protect the domestic kingdom against the foreign enemies. The traditional Muay-Thai of the king had 15 folklores (Thai: แม่ไม้มวยไทย) [4-5] as (1.) Sa Lub Fun Pla, (2.) Pak Sa Waek Rang, (3.) Java Sud Hok, (4.) Tang Krich, (5.) Yok Khao Phra Sumen, (6.) Ta Then Kum Fak, (7.) Morn Yan Lak, (8.) Pak Look Toy, (9.) Jor Ra Kay Fad Harng, (10.) Hak Nguang Ayara, (11.) Na Ka Bid Harng, (12.) Wi Roon Hok Klub, (13.) Dub Chavala, (14.) Koon Yak Jub Ling, and (15.) Huk Khor Erawan, respectively.

Figure 1. "Muay-Thai" – a valuable cultural heritage of martial arts

As of today, Muay-Thai has been tremendously interesting from foreigners coupled with the Thai-tourist destinations [6]. It is not surprise that why many foreigners spend their long holidays to learn Muay-Thai and also travel in Thailand. Significantly, there are many traditional Muay-Thai gyms available for training the foreigners in Bangkok [7] such as Luk Tup Fah Muay Thai Camp (Thai: ค่ายลูกทัพฟ้า), Sor Vorapin Camp (Thai: ค่ายส.วรพิน) and Jiti Gym (Thai: ค่ายมวยไทยจิตติยิม), etc. When many people learn these Muay-Thai folklores, it can be proudly seen as the "inheritance of cultural Thai-ness" that is unique so far than other martial arts. All things considered, it is proven that this present time is the "golden age of Muay-Thai". In order to conserve those 15 Muay-Thai folklores, this paper introduces a "novel thought" as the first groundwork to retrieve the Muay-Thai folklores from the "Muay-Thai action" image dataset. This paper uses image processing [8-9] in conjunction with machine learning [10-11] techniques that are successfully seen in various researches like face detection [12-13], landmark recognition [14-16], tourism recommendation [17-18], cathedral retrieval [19-21], food recognition [22-24], satellite imagery [25-27] or even human action classification [28-30]. As well as the "Muay-Thai folk images" that can be retrieved from the action dataset using their "salient region matching" in terms of displacement magnitudes and angles. The evaluation was done by the average recall and precision that were 0.75 and 0.80, respectively.

This paper is organized into 4 sections. The proposed framework is at section 2. And section 3 and 4 discuss about experimental results and conclusion, respectively.

II. The Muay-Thai Folklore Image Retreival by Salient Region Matching Framework

The framework "Muay-Thai folklore image retrieval by salient region matching" consists of data preparation, object thresholding, image separation, scaling pyramid and region matching as shown in Figure 2.

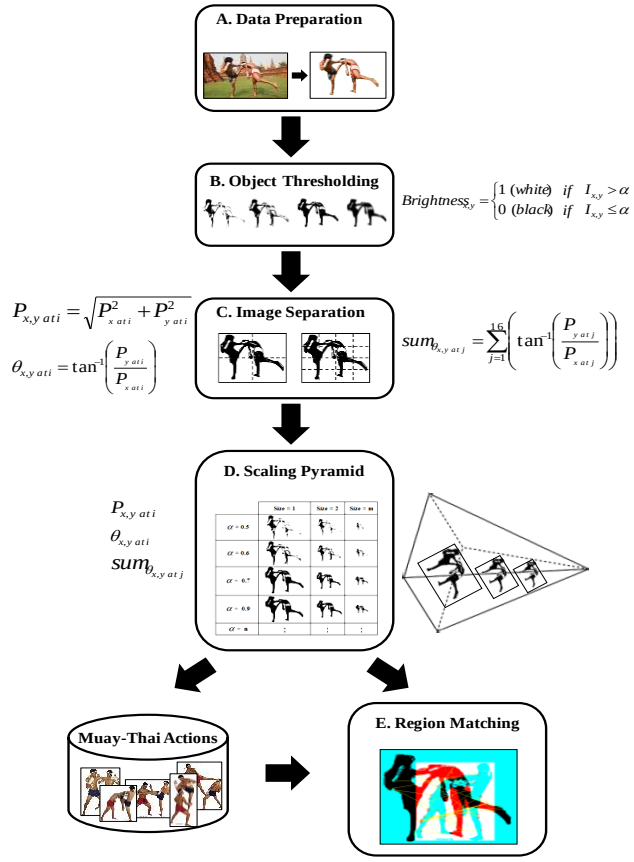


Figure 2. The Muay-Thai folklore image retrieval by salient region matching framework

A. Data Preparation

The 150 Muay-Thai folk lore images with 15 folklore's names are used in our experiment. Both the images for creation of dataset (called "training set") and unknown image (called "testing set") are firstly input to data preparation. All images are also cropped out the noises in terms of background and other unimportant features to remove the noise of recognition.

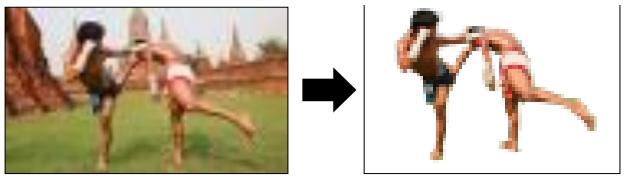


Figure 3. Cropping out of unimportant features.

B. Object Thresholding

The cropped image is set the threshold value into 2 groups as "interested region" and "background" using the brightness threshold (α) filtering by (1).

$$Brightness_{x,y} = \begin{cases} 1\ (white) & if \quad I_{x,y} > \alpha \\ 0\ (black) & if \quad I_{x,y} \le \alpha \end{cases} \tag{1}$$

where $I_{x,y}$ is any pixels within the image

The 4 different thresholds of an image (called binary image) can be shown in Figure 4.

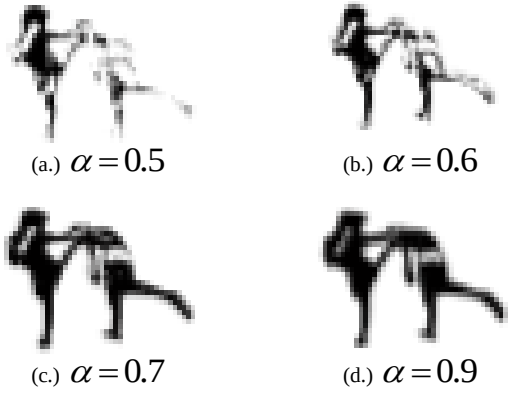
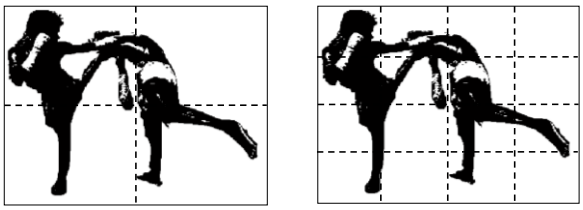


Figure 4. Difference of threshold segmentation

C. Image Seperation

The binary image is seperated into 4 regions (or the 25% of the image's size) as shown in Figure 4. And each region also is seperated into 4^2 = 16 sub-regions as shown in Figure 5.



(a.) Image sepreation (b.) Region seperation

Figure 5. Image and region seperation

Each "i-th" region is computed the magnitude from the "x-th, y-th" pixel values by (2) and the direction of all pixels was done in term of "slope of the regions" by (3).

$$P_{x,y\,at\,i} = \sqrt{P_{x\,at\,i}^2 + P_{y\,at\,i}^2} \tag{2}$$

$$\theta_{x,y\,at\,i} = \tan^{-1}\left(\frac{P_{y\,at\,i}}{P_{x\,at\,i}}\right) \tag{3}$$

Furthermore, each j-th sub-region is computed the orientation summation by (4).

$$sum_{\theta_{x,y\,at\,j}} = \sum_{j=1}^{16}\left(\tan^{-1}\left(\frac{P_{y\,at\,j}}{P_{x\,at\,j}}\right)\right) \tag{4}$$

D. Scaling Pyramid

From the previous data preparation, object thresholding and image separation, all features within a binary image are in terms of "magnitude ($P_{x,y\,at\,i}$)", "direction ($\theta_{x,y\,at\,i}$)" and "orientation summation ($sum_{\theta_{x,y\,at\,j}}$)", respectively. The binary image is also scaled down and also repeatedly computed those 3 terms of features in the m different sizes, called the "scaling pyramid", as shown in Figure 6.

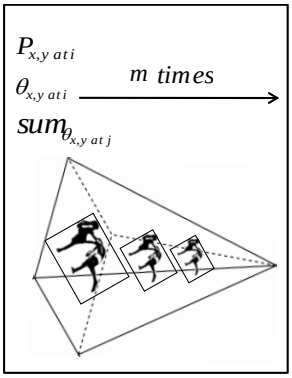


Figure 6. Scaling pyramid in different sizes

Not only the scaling pyramid is in different "m sizes" but also it is done in the "n thresholds". The output of pyramid is a "vector representation of image". The vector keeps all features within the image in term of floating numbers (also called "feature description"). To that end, the complexity of all scaling pyramids can be iterative as " $O(mn)$ ", which is shown in Figure 7.

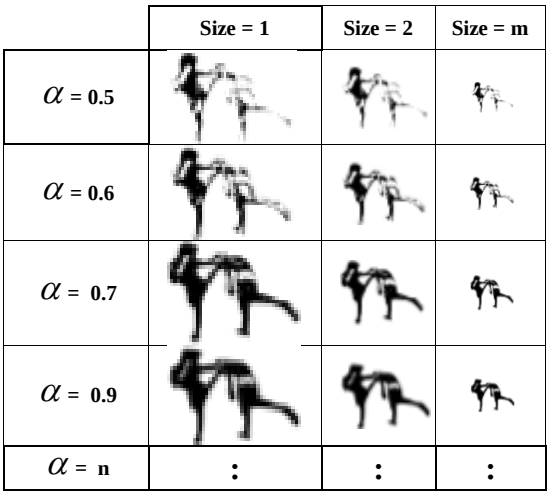
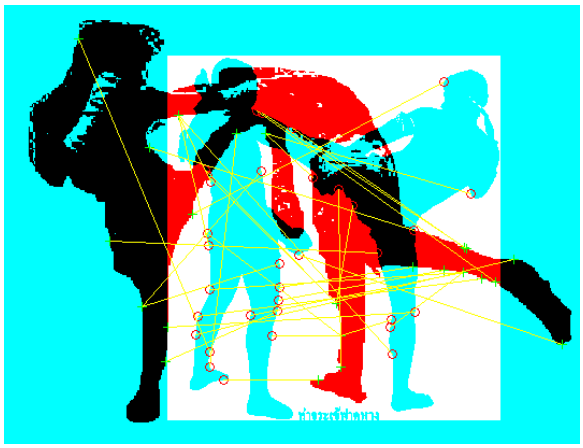


Figure 7. Scaling pyramid in different sizes and threshold

E. Region Matching

For training set, all Muay-Thai folklore images in term of vectors with their textual information are kept in the "Muay-Thai action dataset" that will be used to match with the unknown folklore image. In contrast, the region matching is a process to compare the unknown image (in term of vectors) within the test set and the vectors (with their information) from the training sets. In comparison, the most similar image from the dataset is selected to match with the unknown image. And the information annotation is also done, as shown in Figure 8.



Jor Ra Kay Fad Harng :
...Protected the punch by kicking
This technique uses the heel reverse kick...

Figure 8. The region matching between the "unknown image" and the "most similar image" from the action dataset

The similarity of images can be measured using Euclidean distance. The most similar image from the dataset must have the lowest Euclidean distance value because it has the shortest distance (or the nearest image) compared to the unknown image. The Euclidean distance can be computed by (5).

$$Dist_{Euclidean} = \sqrt{\sum_{1}^{n}\left(I_{dataset} - I_{unknown}\right)^2} \quad (5)$$

III. EXPERIMENTAL RESULTS

This experiment used 150 Muay-Thai folklore images that covered 15 folklores as (1.) Sa Lub Fun Pla, (2.) Pak Sa Waek Rang, (3.) Java Sud Hok, (4.) Tang Krich, (5.) Yok Khao Phra Sumen, (6.) Ta Then Kum Fak, (7.) Morn Yan Lak, (8.) Pak Look Toy, (9.) Jor Ra Kay Fad Harng, (10.) Hak Nguang Ayara, (11.) Na Ka Bid Harng, (12.) Wi Roon Hok Klub, (13.) Dub Chavala, (14.) Koon Yak Jub Ling, and (15.) Huk Khor Erawan, respectively.

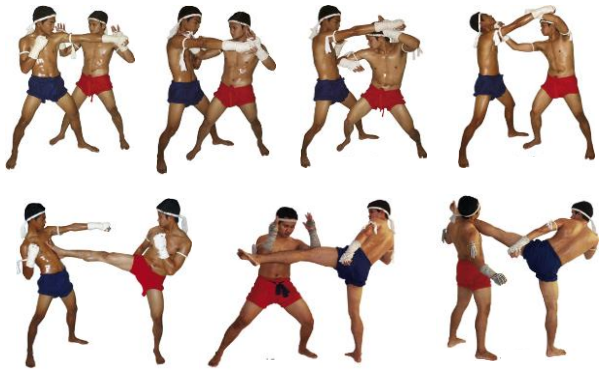


Figure 9. Some Muay-Thai folklores

We have done our experiment on top of the content-based image retrieval using Matlab 2016a. Our proposed framework implementation was based on "Harris corner detection" [31] coupled with "Dense-SIFT description" [32]. In light of measurement, we used the "recall" and "precision" to evaluate our experiment by (6) and (7).

$$Recall = \frac{TP}{TP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

- TP (True Positive) is "the folklore A" and the system correctly detects as "the folklore A"
- FP (False Positive) is "not the folklore A" and the system wrongly detects as "the folklore A"
- FN (False Negative) is "the folklore A" and the system wrongly detects as "not the folklore A"

Actual Muay-Thai Folklore \ Predicted Muay-Thai Folklore	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	7		1	2											
2		6	4												
3		1	8		1										
4	1			8									1		
5	1				8	1									
6					1	9									
7							7	2		1					
8								6		4					
9									10						
10								5		5					
11											7	3			
12												10			
13				1									9		
14	1			2								2		5	
15										2					8

Figure 10. The experimental results between actual and predicted Muay-Thai folklores

The overall experimental results are shown in Figure 10. The recall and precision of all 15 folklores are shown in TABLE I.

TABLE I. THE RECALL AND PRECISION OF 15 FOLKLORES

Muay-Thai Folklore	Recall	Precision
1. Sa Lub Fun Pla	0.7	0.7
2. Pak Sa Waek Rang	0.6	0.86
3. Java Sud Hok	0.8	0.62
4. Tang Krich	0.8	0.62
5. Yok Khao Phra Sumen	0.8	0.8
6. Ta Then Kum Fak	0.9	0.9
7. Morn Yan Lak	0.7	1
8. Pak Look Toy	0.6	0.46
9. Jor Ra Kay Fad Harng	1	1
10. Hak Nguang Ayara	0.5	0.42
11. Na Ka Bid Harng	0.7	1
12. Wi Roon Hok Klub	1	0.67
13. Dub Chavala	0.9	0.9
14. Koon Yak Jub Ling	0.5	1
15. Huk Khor Erawan	0.8	1
Average	**0.75**	**0.8**

From the experimental results, the "Jor Ra Kay Fad Harng (Thai: จระเข้ฟาดหาง)" folklore had the 100% correctness in term of recall and precision. Since this folklore was totally distinctive from other folklores. Having said that, "Koon Yak Jub Ling (Thai: ขุนยักษ์จับลิง)" was the most complicated Muay-Thai folklore that had the low recall as 0.5. Moreover, the "Pak Look Toy (Thai: ปักลูกทอย)" and "Hak Nguang Ayara (Thai: หักงวงไอยรา)" sometimes were looked almost similar (as shown in Figure 11) that made their correctness were lower than other folklores. With this in mind, the average recall and precision were 0.75 and 0.8, respectively

(a.) Pak Look Toy (b.) Hak Nguang Ayara

Figure 11. The similarity between "Pak Look Toy" and "Hak Nguang Ayara"

IV. CONCLUSION

Since Muay-Thai has become one of the best martial arts in the world that is called the golden age of Muay-Thai. The art of eight limbs (Thai: นวอาวุธ) in Muay-Thai consists of 2-fists, 2-elbows, 2-knees, and 2-shins to fight the opponent that makes Muay-Thai is distinguish from other martial arts. Muay-Thai was officially disseminated by the king Suriyentharathibodhi of Ayutthaya. The original Muay-Thai consisted of 15 folklores as (1.) Sa Lub Fun Pla, (2.) Pak Sa Waek Rang, (3.) Java Sud Hok, (4.) Tang Krich, (5.) Yok Khao Phra Sumen, (6.) Ta Then Kum Fak, (7.) Morn Yan Lak, (8.) Pak Look Toy, (9.) Jor Ra Kay Fad Harng, (10.) Hak Nguang Ayara, (11.) Na Ka Bid Harng, (12.) Wi Roon Hok Klub, (13.) Dub Chavala, (14.) Koon Yak Jub Ling, and (15.) Huk Khor Erawan, respectively. Thereupon, Muay-Thai is the valuable cultural heritage of Thailand and Thai signature. This paper is proposes a "novel thought" in the combination of "Muay-Thai folklores" and "human action classification" to conserve

"Thai-ness (Thai: ความเป็นไทย)" and "Thai Culture (Thai: วัฒนธรรมไทย)" using image processing coupled with machine learning. This experiment was evaluated by average "recall" and "precision" that were summarized as 0.75 and 0.8, respectively.

REFERENCES

[1] "9 Reasons Why Muay Thai Is The Perfect Martial Art," [Online]. Available: https://evolve-mma.com/blog/9-reasons-muay-thai-perfect-martial-art/. [Accessed: 24 February 2018].

[2] "Preserving The Classic Technique of Maemai Muay Thai," [Online]. Available: http://fightland.vice.com/blog/preserving-the-classic-technique-of-maemai-muay-thai. [Accessed: 25 February 2018].

[3] Department of Cultural Promotion, "Muay-Thai as the pride and heritage of Thai people from the ancestors," [Online]. Available: http://www.culture.go.th/culture_th/mobile_detail.php?cid=11&nid=584. [Accessed: 25 February 2018]. (in Thai).

[4] Thai Encyclopedia for juveniles, "The 15 Muay-Thai Folklores," [Online]. Available: http://kanchanapisek.or.th/kp6/sub/book/book.php?book=35&chap=3&page=t35-3-infodetail05.html. [Accessed: 26 February 2018]. (in Thai).

[5] Thailandee, "Ceremony of World Wai Kru Muay Thai," [Online]. Available: https://www.thailandee.com/en/events-thailand/ceremony-of-world-wai-kru-muay-thai-62. [Accessed: 26 February 2018].

[6] "9 Reasons You Need To Train Muay Thai in Thailand," [Online]. Available: http://muay-thai-guy.com/train-muay-thai-in-thailand.html. [Accessed: 26 February 2018].

[7] Bangkok.com, "10 Great Muay Thai Gyms in Bangkok," [Online]. Available: http://www.bangkok.com/magazine/best-muay-thai-gyms.htm. [Accessed: 26 February 2018].

[8] A. A. Olaode, G. Naghdy and C. A. Todd, "Unsupervised Image Classification by Probabilistic Latent Semantic Analysis for the Annotation of Images," 2014 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Wollongong, NSW, 2014, pp. 1-8.

[9] P. Mookdarsanit, L. Soimart, M. Ketcham and N. Hnoohom, "Detecting Image Forgery Using XOR and Determinant of Pixels for Image Forensics," 2015 11th International Conference on Signal-Image Technology & Internet-Based Systems, Bangkok, 2015, pp. 613-616.

[10] A. Ahar, A. Barri and P. Schelkens, "From Sparse Coding Significance to Perceptual Quality: A New Approach for Image Quality Assessment," in IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 879-893, Feb. 2018.

[11] L. Soimart and P. Pongcharoen, "Multi-row Machine Layout Design using Aritificial Bee Colony," 2011 International Conference on Economics and Business Information, Bangkok, 2011.

[12] L. Soimart and P. Mookdarsanit, "Gender Estimation of a Portrait: Asian Facial-significance Framework," The 6th International Conference on Sciences and Social Sciences, Mahasarakham, 2016.

[13] Y. Luo and Y. P. Guan, "Adaptive skin detection using face location and facial structure estimation," in IET Computer Vision, vol. 11, no. 7, pp. 550-559, 10 2017.

[14] X. Qian et al., "Image Location Inference by Multisaliency Enhancement," in IEEE Transactions on Multimedia, vol. 19, no. 4, pp. 813-821, April 2017.

[15] P. Mookdarsanit and M. Rattanasiriwongwut, "Location Estimation of a Photo: A geo-signature MapReduce Workflow," in Engineering Journal, vol.21, no.3, pp.295-308, May 2017.

[16] P. Mookdarsanit and M. Ketcham, "Image Location Estimation of Well-known Places from Multi-source based Information," The 11th International Symposium on Natural Language Processing, Ayutthaya, 2016.

[17] L. F. D'Haro, R. E. Banchs, C. K. Leong, L. G. M. Daven and N. T. Yuan, "Automatic labelling of touristic pictures using CNNs and metadata information," 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), Singapore, 2017, pp. 292-296.

[18] L. Soimart and P. Mookdarsanit, "Name with GPS Auto-tagging of Thai-tourist Attractions from An Image," The 2017 Technology Innovation Management and Engineering Science International Conference, 2017.

[19] P. Mookdarsanit and S. Rattanasiriwongwut, "GPS Determination of Thai-temple Arts from a Single Photo," The 11th International Conference on on Applied Computer Technology and Information Systems, Bangkok, Thailand, 2017, pp. 42-47.

[20] J. Lv, X. Shao, J. Xing, C. Cheng and X. Zhou, "A Deep Regression Architecture with Two-Stage Re-initialization for High Performance Facial Landmark Detection," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 3691-3700.

[21] P. Mookdarsanit and M. Rattanasiriwongwut, "MONTEAN Framework: A Magnificent Outstanding Native-Thai and Ecclesiastical Art Network," in International Journal of Applied Computer Technology and Information Systems, vol.6, no.2, pp.17-22, 2017.

[22] H. He, F. Kong and J. Tan, "DietCam: Multiview Food Recognition Using a Multikernel SVM," in IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 3, pp. 848-855, May 2016.

[23] L. Soimart and P. Mookdarsanit, "Ingredients estimation and recommendation of Thai-foods," in SNRU Journal of Science and Technology, vol.9, no.2, pp.509-520, Jul. 2017.

[24] G. Ciocca, P. Napoletano and R. Schettini, "Food Recognition: A New Dataset, Experiments, and Results," in IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 3, pp. 588-598, May 2017.

[25] L. Soimart and M. Ketcham, "An efficient algorithm for earth surface interpretation from satellite imagery," in Engineering Journal, vol.20, no.5, pp.215-228, Nov. 2016.

[26] L. Soimart and M. Ketcham, "Hybrid of Pixel-based and Region-based Segmentation for Geology Exploration from Multi-spectral Remote Sensing," The 11th International Symposium on Natural Language Processing, Ayutthaya, 2016.

[27] L. Soimart and M. Ketcham, "The Segmentation of Satellite Image Using Transport Mean-shift Algorithm," 13th International Conference on IT Applications and Management (ITAM-13), 2015, pp. 124-128.

[28] Z. Zhao, H. Ma and S. You, "Single Image Action Recognition Using Semantic Body Part Actions," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 3411-3419.

[29] P. Mookdarsanit and L. Mookdarsanit, "An Automatic Image Tagging of Thai Dance's Gestures," Joint Conference on ACTIS & NCOBA, Ayutthaya, Thailand, January 2018, pp. 76-80.

[30] Y. Lavinia, H. H. Vo and A. Verma, "Fusion Based Deep CNN for Improved Large-Scale Image Action Recognition," 2016 IEEE International Symposium on Multimedia (ISM), San Jose, CA, 2016, pp. 609-614.

[31] B. A. Jasani, S. K. Lam, P. K. Meher and M. Wu, "Threshold-Guided Design and Optimization for Harris Corner Detector Architecture," in IEEE Transactions on Circuits and Systems for Video Technology, vol. PP, no. 99, pp. 1-1.

[32] Q. Zhou, u. R. Shafiq, Y. Zhou, X. Wei, L. Wang and B. Zheng, "Face Recognition Using Dense SIFT Feature Alignment," in Chinese Journal of Electronics, vol. 25, no. 6, pp. 1034-1039, 11 2016.