RGC Ref. No.:
UGC/FDS14/E02/14
(please insert ref. above)

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report
(for completed projects only)

Submission Deadlines:
1. *Auditor's report with unspent balance, if any: within **six** months of the approved project completion date.*
2. *Completion report: within **12** months of the approved project completion date.*

Part A: The Project and Investigator(s)

1. Project Title

Model-based Unsupervised Image Segmentation

2. Investigator(s) And Academic Department(s) / Unit(s) Involved

Research Team	Name / Post	Unit / Department / Institution
Principal Investigator	CHOY Siu-kai / Associate Professor	Department of Mathematics and Statistics / Hang Seng Management College
Co-Investigator(s)	LAM Benson Shu-yan / Assistant Professor	Department of Mathematics and Statistics / Hang Seng Management College
Others		

3. Project Duration

	Original	Revised	Date of RGC / Institution Approval *(must be quoted)*
Project Start Date	01 Jan 2015	01 Jan 2015	NA
Project Completion Date	31 Dec 2016	30 Jun 2017	12 Sep 2016
Duration *(in month)*	24	30	12 Sep 2016
Deadline for Submission of Completion Report	31 Dec 2017	30 Jun 2018	12 Sep 2016

Part B: The Final Report

5. Project Objectives

5.1 Objectives as per original application

1. To develop a refined bit-plane-based probability model for parametric random histogram modeling and justify its use by studying the statistical properties of the model.
2. To investigate an agglomerative fuzzy *k*-means algorithm with spatial information for various data-mining purposes.
3. To develop a mathematical optimization framework to integrate the refined bit-plane-based probability model with the agglomerative fuzzy *k*-means algorithm with spatial information.
4. To investigate the feasibility of the proposed model-based clustering algorithm and the mathematical theory that underpins it.
5. To develop a robust and efficient unsupervised image segmentation algorithm.
6. To conduct extensive experimental testing of the proposed method in various image segmentation applications.
7. To extend the applicability of the proposed framework to additional fields, such as intelligent video surveillance.

5.2 Revised objectives

Date of approval from the RGC:	09 December 2014
Reasons for the change:	The changes of the project objectives are due to the reduction of budget requested and project duration.

1. To develop a refined bit-plane-based probability model for parametric random histogram modeling and justify its use by studying the statistical properties of the model.
2. To investigate an agglomerative fuzzy *k*-means algorithm with spatial information for various data-mining purposes.
3. To develop a mathematical optimization framework to integrate the refined bit-plane-based probability model with the agglomerative fuzzy *k*-means algorithm with spatial information.
4. To develop a robust unsupervised image segmentation algorithm and conduct extensive experimental testing of the proposed method in various image segmentation applications.

5.3 Realisation of the objectives

(Maximum 1 page; please state how and to what extent the project objectives have been achieved; give reasons for under-achievements and outline attempts to overcome problems, if any)

All four project objectives have been achieved.

Objective 1: To develop a refined bit-plane-based probability model for parametric random histogram modeling and justify its use by studying the statistical properties of the model.

- We have developed a bit-plane-dependence probability model, which is an improved version of product of Bernoulli distributions, for characterizing image variations that do not need to have specific structure. Specifically, the proposed model provides a universal parametric representation that can be used to model random histograms without enforcing any specific restrictions on the histograms.
- We have conducted statistical analysis on the proposed model and obtained the closed-form expression of maximum likelihood (ML) estimators of model parameters. In addition, we show that the ML estimators are joint sufficient statistics, which, in turn, justify the theoretical basis for their use.

Objectives 2&3: To investigate an agglomerative fuzzy *k*-means algorithm with spatial information for various data-mining purposes; and to develop a mathematical optimization framework to integrate the refined bit-plane-based probability model with the agglomerative fuzzy *k*-means algorithm with spatial information.

- We have developed a fuzzy bit-plane-dependence image segmentation model, which integrates the bit-plane-dependence probability model and spatial/neighboring information into the agglomerative fuzzy *k*-means algorithm. We have also developed an optimization procedure to minimize the fuzzy bit-plane-dependence image segmentation model based on the alternating minimization procedure.

Objective 4: To develop a robust unsupervised image segmentation algorithm and conduct extensive experimental testing of the proposed method in various image segmentation applications.

- We have developed a fuzzy bit-plane-dependence unsupervised image segmentation algorithm and conducted extensive experiments on different kinds of images using public databases. Comparative experimental results with existing fuzzy unsupervised image segmentation approaches show that our method achieves remarkable success in image segmentation applications.

5.4 Summary of objectives addressed to date

Objectives *(as per 5.1/5.2 above)*	**Addressed** *(please tick)*	**Percentage Achieved** *(please estimate)*
1. To develop a refined bit-plane-based probability model for parametric random histogram modeling and justify its use by studying the statistical properties of the model.	✓	100%
2. To investigate an agglomerative fuzzy *k*-means algorithm with spatial information for various data-mining purposes.	✓	100%
3. To develop a mathematical optimization framework to integrate the refined bit-plane-based probability model with the agglomerative fuzzy *k*-means algorithm with spatial information.	✓	100%
4. To develop a robust unsupervised image segmentation algorithm and conduct extensive experimental testing of the proposed method in various image segmentation applications.	✓	100%

6. Research Outcome

6.1 Major findings and research outcome
(Maximum 1 page; please make reference to Part C where necessary)

The major findings and research outcome, which were published or submitted for publication as in Part C, are shown below:

1. We have proposed a fuzzy Generalized Gaussian Density (GGD) segmentation model and the GGD-based agglomerative fuzzy algorithm for clustering image pixels. The merits of algorithm are that it is not sensitive to initial parameters and that the number of clusters can be estimated via the validation technique. To minimize the objective function of the model, we define a dissimilarity measure based on the Kullback–Leibler divergence of the GGDs that computes the discrepancy between GGDs in the space of generalized probability distributions. To effectively segment images with various textures, we propose a two-stage fuzzy GGD segmentation algorithm. The first stage adopts the proposed fuzzy algorithm to obtain initial segmentation and the second stage improves initial segmentation by image boundary correction. Experimental results show that our proposed method has a promising performance compared with existing approaches.

2. We have proposed a probability model for characterizing the distributions of image variations based on bit-plane probabilities and dependencies between bit-planes. Compared with the current state-of-the-art image variation models which assume the distributions have specific structures (e.g., symmetry, monotone and periodicity), the proposed model provides a universal parametric representation that can be used to model random distributions without enforcing any specific restrictions on the distributions. In addition, we show that the maximum likelihood estimators of model parameters are joint sufficient statistics, which, in turn, justify the theoretical basis for their use. To effectively segment images with various textures, we propose a fuzzy bit-plane-dependence image segmentation algorithm. The proposed algorithm integrates the bit-plane-dependence probability model into the agglomerative fuzzy algorithm, and incorporates spatial/neighboring information and boundary localization for image segmentation applications. Experiments demonstrate the superior performance of the proposed method.

6.2 Potential for further development of the research and the proposed course of action *(Maximum half a page)*

The project can be further developed by studying the following aspects:

1. A better segmentation result is expected if both spatial and frequency image features are incorporated into the segmentation model. Such spatial/frequency features (together with the appropriate dissimilarity measures for different features) can be easily incorporated into our framework by modifying the proposed objective function.
2. Statistical analysis on the intrinsic structures of bit-planes should be performed in order to have a comprehensive understanding of the dependencies between bit-planes. The results are expected to provide statistical justification on the existence of bit-plane dependencies and the selection of the order of bit-plane dependencies.
3. The relationship between model parameters (i.e., bit-plane probabilities and bit-plane dependencies) and real-world instance should be studied. Specifically, we should study the physical meaning of model parameters and "how do the model parameters relate to the objects that are present in the image".

7. Layman's Summary

(Describe in layman's language the nature, significance and value of the research project, in no more than 200 words)

Clustering algorithms and raw histograms have a wide range of applications in the fields of image processing and data mining. However, traditional clustering algorithms are sensitive to initial clusters, and the number of clusters is unknown a priori while raw histograms are high-dimensional and are thus less efficient for applications. In response to these difficulties, the goal of the proposed project is to investigate a robust and efficient clustering algorithm and develop an effective parametric random histogram model combined with a boundary localization technique for image segmentation applications. More specifically, we will investigate an agglomerative fuzzy *k*-means algorithm with spatial information and incorporate a refined bit-plane-based probability model into the clustering process. The proposed method is expected to remedy the main drawbacks of both traditional/modern clustering algorithms and raw histograms simultaneously, and thus is expected to achieve promising image segmentation results. The combination of our proposed clustering algorithm and random histogram modelling technique may thus assist researchers in the fields of data mining and statistical learning by providing a new mode-based unsupervised learning algorithm with practical and real-time computer vision applications.

Part C: Research Output

8. Peer-Reviewed Journal Publication(s) Arising Directly From This Research Project

(Please attach a copy of the publication and/or the letter of acceptance if not yet submitted in the previous progress report(s). All listed publications must acknowledge RGC's funding support by quoting the specific grant reference.)

The Latest Status of Publications				Author(s) *(denote the corresponding author with an asterisk*)*	Title and Journal / Book *(with the volume, pages and other necessary publishing details specified)*	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this Report *(Yes or No)*	Acknowledged the Support of RGC *(Yes or No)*	Accessible from the institutional repository *(Yes or No)*
Year of Publication	**Year of Acceptance** *(For paper accepted but not yet published)*	**Under Review**	**Under Preparation** *(optional)*						
2017				Choy Siu Kai*, Lam Shu Yan, Yu Kwok Wai, Lee Wing Yan and Leung King Tai	Fuzzy Model-based Clustering and Its Application in Image Segmentation (vol.68, pp. 141-157, Pattern Recognition)	**No**	Yes	Yes	Yes
2019	2018			Choy Siu Kai*, Yuen Fei Lung and Yu Kwok Wai,	Fuzzy Bit-plane-dependence Image Segmentation. Signal Processing, Vol. 154, pp. 30-44.	**No**	Yes	Yes	Yes

9. Recognized International Conference(s) In Which Paper(s) Related To This Research Project Was / Were Delivered

(Please attach a copy of each conference abstract)

Month / Year / Place	Title	Conference Name	Submitted to RGC *(indicate the year ending of the relevant progress report)*	Attached to this Report *(Yes or No)*	Acknowledged the Support of RGC *(Yes or No)*	Accessible from the institutional repository *(Yes or No)*
April / 2015 / South Brisbane	K-medians clustering based ℓ1-PCA model	IEEE International Conference on Acoustics, Speech and Signal Processing	2015	Yes	Yes	Yes

10. Whether Research Experience And New Knowledge Has Been Transferred / Has Contributed To Teaching And Learning

(Please elaborate)

The proposed algorithm can be used as examples in some courses related to data mining and machine learning.

11. Student(s) Trained

(Please attach a copy of the title page of the thesis)

Name	Degree Registered for	Date of Registration	Date of Thesis Submission / Graduation
NA			

12. Other Impact

(e.g. award of patents or prizes, collaboration with other research institutions, technology transfer, teaching enhancement, etc.)

NA

13. Public Access Of Completion Report

(Please specify the information, if any, that cannot be provided for public access and give the reasons.)

Information that Cannot Be Provided for Public Access	Reasons
NA	

RESEARCH GRANTS COUNCIL
COMPETITIVE RESEARCH FUNDING SCHEMES FOR
THE LOCAL SELF-FINANCING DEGREE SECTOR

FACULTY DEVELOPMENT SCHEME (FDS)

Completion Report - Attachment
(for completed projects only)

RGC Ref. No.:	UGC/FDS14/E02/14
Principal Investigator:	CHOY Siu-kai
Project Title:	Model-based Unsupervised Image Segmentation

Statistics on Research Outputs

	Peer-reviewed Journal Publications	Conference Papers	Scholarly Books, Monographs and Chapters	Patents Awarded	Other Research Outputs (Please specify)
No. of outputs arising directly from this research project [or conference]	2	0	0	0	0