APPLYING LCS TO AFFECTIVE IMAGE CLASSIFICATION IN SPATIAL-FREQUENCY DOMAIN

Po-Ming Lee[1] and Tzu-Chien Hsiao[2]

[1]*Institute of Computer Science and Engineering, Department of Computer Science, National Chiao Tung Universityal. 1001 University Rd., Hsinchu, Taiwan, R.O.C.*

[2]*Department of Computer Science, Institute of Biomedical Engineering, and Biomedical Electronics Translational Research Center and Biomimetic Systems Research Center in National Chiao Tung Universityal. 1001 University Rd., Hsinchu, Taiwan, R.O.C.*

Abstract

Recent studies have utilizes color, texture, and composition information of images to achieve affective image classification. However, the features related to spatial-frequency domain that were proven to be useful for traditional pattern recognition have not been tested in this field yet. Furthermore, the experiments conducted by previous studies are not internationally-comparable due to the experimental paradigm adopted. In addition, contributed by recent advances in methodology, that are, Hilbert-Huang Transform (HHT) (i.e. Empirical Mode Decomposition (EMD) and Hilbert Transform (HT)), the resolution of frequency analysis has been improved. Hence, the goal of this research is to achieve the affective image-classification task by adopting a standard experimental paradigm introduces by psychologists in order to produce international-comparable and reproducible results; and also to explore the affective hidden patterns of images in the spatial-frequency domain. To accomplish these goals, multiple human-subject experiments were conducted in laboratory. Extended Classifier Systems (XCSs) was used for model building because the XCS has been applied to a wide range of classification tasks and proved to be competitive in pattern recognition. To exploit the information in the spatial-frequency domain, the traditional EMD has been extended to a two-dimensional version. To summarize, the model built by using the XCS achieves Area Under Curve (AUC) = 0.91 and accuracy rate over 86%. The result of the XCS was compared with other traditional machine-learning algorithms (e.g., Radial-Basis Function Network (RBF Network)) that are normally used for classification tasks. Contributed by proper selection of features for model building, user-independent findings were obtained. For example, it is found that the horizontal visual stimulations contribute more to the emotion elicitation than the vertical visual stimulation. The effect of hue, saturation, and brightness; is also presented.

1 Introduction

1.1 Scope

People experience emotion in their daily life by feeling happy, angry and various emotions induced by stimulus and events that are emotionally relevant. Because it is human nature to pursue happiness and avoid pain, a research finding related to human emotion can be easily transferred to diverse applications. For example, behavioral economics [1], media studies and advertisement [2, 3]. Some researches focused on the use of emotional relevant stimulus to attract the attention of subjects, and to make subjects remember more on the product presented [3]. In the area of image, print advertise-

ment and the use of affective images for attracting the attention of subjects during web browsing were reported [2]. Guideline of extracting emotional relevant features in a web page is also available [4].

Due to the development of personal computer, software, and World-Wide-Web (WWW), people nowadays generate huge amount of content (e.g., daily news, articles on variety topics and personal data) and upload them to the internet every day. To enable end users to explore the content on the internet, Google and Yahoo! such the search engine provider index these contents. Currently most of the web content indexing works are done based on text-based technologies. Although text-based indexing technologies are suitable for articles, the limitation of the text-based method is obvious when images are the indexing target. Traditionally, image search is done based on the file name of the target image and perhaps the description (e.g., tags) of the target image. In the last decade, image search based on content has been provided by Google Picture and Yahoo! Image Search. However, there was lack of attention to the development of the techniques of indexing the affective characteristics of images, despite equipped with such the technique, the industry would be able to design new applications related to human feelings and better user experience. For example, an application that leads end users to target images that may potentially ease their "feelings".

1.2   Motivation

The affective characteristic of an image is defined by the capability of an image in eliciting emotional responses. Human beings have the ability to recognize the affective characteristic embedded in an image. Hence, to index the affective characteristics of images on the internet, an intuitive approach is to have a large number of people manually rate all the images and calculate descriptive statistics from the obtained ratings. However, this approach may be impractical due to the cost of manpower and the increasing speed of images around the world.

On the other hand, using Artificial Intelligence (AI) and Machine Learning (ML) techniques, a broad range of intelligent machines have been designed to perform different pattern recognition tasks [5-7]. An intelligent machine that can automatically classify images based on their affective characteristics could be built by given a number of instances

with proper selected features. Due to the lack of attention on this issue in the literature, this study aims to build an intelligent machine to perform affective image-classification task.

1.3   Research Objectives

he proposed hypothesis is that a trained intelligent machine can classify images based on their ability in eliciting emotions, through the basic properties of these images. Wilson's Extended Classifier System (XCS) [8], a well-tested accuracy-based Learning Classifier System (LCS) model, is to be used to build the classification models in this research. The XCS is proven to be capable of extracting complete, general, and readable rules from a previously unknown dataset, which motivated its suitability for this research work.

The overall goal of this research is to demonstrate a novel method to classify images based on their ability in eliciting emotions. This goal is divided into the following two subgoals.

– To develop an intelligent machine that can identify images based on their capability in inducing emotions. To examine the effect of basic properties (i.e. hue, saturation, and brightness) of images on their capability in inducing emotions.

– To develop an intelligent machine that can identify images based on their capability in inducing emotions. To examine the effect of the properties of images in the spatial-frequency domain on the capability of these images in inducing emotions.

All the human-subject experiments conducted in this research, and the manner of using data obtained from human subjects were approved (Protocol No: 100-014-E and NCTU-REC-102-007) by the Institution Review Board (IRB) of the National Taiwan University Hospital Hsinchu Branch and the IRB of National Chiao-Tung University, respectively.

The built models were evaluated using 10-Fold Cross Validation (CV) which is a traditional evaluation method used in the literature and the results were compared with the existing related systems. In addition to the demonstration of the modeling building process, this study also aims at providing

the examination results of the factors that may influence the affective characteristics of images.

1.4 Paper Contribution

This research led to the following major contributions to the field of affective computing [9] in general and specifically to the field of affective image classification.

This resarch:(1) Demonstrates a two-dimensional version of the Hilbert-Huang Transform (2D-HHT) for extracting features in the spatial-frequency domain from images. (2) Demonstrates the model-building process of an intelligent machine for performing affective image-classification task. (3) Examines the influence of basic properties and also the properties in regard to the spatial-frequency domain on the affective characteristics of images. (4) All the results were obtained from and validated by human-subject experiments.

1.5 Paper Organization

The remainder of this research is organized as follows. Chapter 2 describes the research methodology to be used in this work to achieve the overall goal. This chapter describes the framework, the emotional stimuli, and the instruments to be used for model building. Chapter 3 and chapter 4 present major contributions to fulfill the established research objectives. Chapter 5 concludes this work.

Chapter 2 describes the research paradigm adopted in this work to achieve the overall goal, and briefly describes the instruments used. Chapter 2 also provides a detailed description of the XCSs along with an overview of related studies.

The details of each literature review of related work and the implemented systems are provided in a separate contribution chapter, that are Chapter 3 and Chapter 4. These two chapters also provide details of the problem domains experimented here and the experimental setup used for collecting data set or testing and evaluating the developed systems. Chapter 3 and 4 describe the two affect detectors that are successfully built from human-subject experiments.

Chapter 5 presents the achieved objectives, main conclusions from each contribution chapter,

and the future work that stems from this research work.

2 Method and Materials

2.1 Emotion Theories

One of the difficulties in studying emotion is that how to define it. Although there is a tendency for researchers to intuitively define a set of discrete basic emotions (e.g., happy, surprising, sad, and angry [10]), recently the dimensional theory of emotion, in replacement of the traditional assumption of discrete emotions, has been proposed and demonstrated to be more suitable than the traditional manner of describing emotions in a number of studies [11, 12].

Dimensional theory defines emotions by a two dimensional affective space, of which the two dimensions are "valence" and "arousal". The valence represents that whether the emotion experienced is pleasant whereas the arousal represents the amplitude of the emotion aroused. The philosophy of the theory adopted is illustrated in Figure 1.
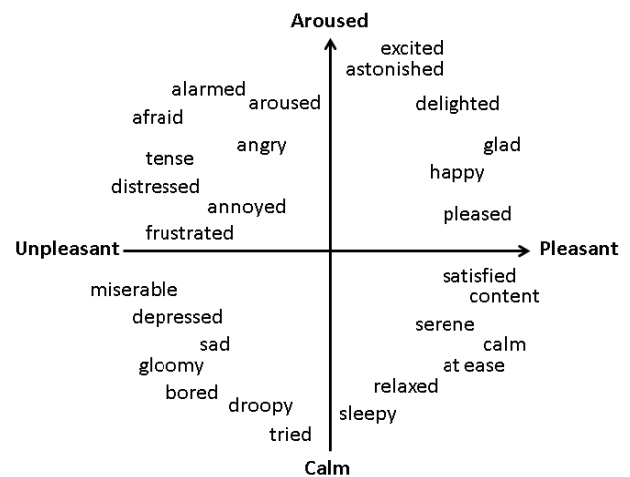

Figure 1. Definition of emotions in a two-dimension affective space

The dimensional theory of emotion explains how human emotion is elicited and the roles of emotional stimulus plays in emotion elicitation by relating the emotion theory to the motivational system of human. The motivational system guides human to behave in the tendency of "approach" or "avoidance" when presented with emotionally relevant stimulus (the "stimulus" can be an object, a

scenario, or a type of circumstance)[13]. The reason of a stimulus to be emotionally relevant could be considered as a result of evolutionary process, that is, can be related to the need of survival. For example, the stimulus that stimulates positive emotions is found related to food and sex, whereas the stimulus that stimulates negative emotions was found related to danger and death. The umbrella term "emotionally relevant" can simply be understood as a capability to elicit certain emotions of a person (either positive or negative emotions) [14].

The dimensional theory of emotion has attracted substantial attention in the field of psychology since proposed, and is commonly adopted in latest studies [3, 11, 15]. On the other hand, brain scientists focused on biological proof. The pathway, the mechanisms of brain, the autonomic nervous system, and the organs, that are accounted for emotion responses have being revealed [16]. Other research reported the experimental results on the relationship between emotion and decision making [17], and also the relationship between emotion and memory [18].

On the path of research in affective image classification [19-23], the achieved accuracy rates are relatively low. Furthermore, the use of definition in "discrete" emotions also caused these experiments hard to reproduce in countries other than the United States. Hence, this study adopts the paradigm of the work related to the dimensional theory of emotion.

2.2   Instruments

2.3   International Affective Picture System (IAPS)

Image is a type of visual stimulus that commonly used in human emotion study for emotion induction. However, in past decades, due to cultural difference, results obtained from different experiments were incomparable. Subsequently, a standard affective picture system named International Affective Picture System (IAPS) was proposed [24] to help emotion research experimenter in providing comparable experimental results.

The IAPS database is developed and distributed by the NIMH Center for Emotion and Attention (CSEA) at the University of Florida to provide a set of normative emotional stimuli for experimental investigations of emotion and attention and can be easily obtained through e-mail application. The IAPS contains various affective pictures selected based on the statistics obtained from experimental results. These pictures are proved to be capable in inducing diverse emotions in the affective space [12]. The IAPS also describes a protocol that includes the constraint about the number of images used in a single experiment and the distribution of the emotions induced by the images selected.

The IAPS has attracted attention since proposal; various experiments, for example, the empirical studies on psychophysiological signals that are related to emotional responses [25], the experiments of the effects of emotion on memory [26], and the experiments for identifying the relationship between motivation and emotion [12], were conducted using IAPS for emotion induction. The images used in this research were solely selected from this IAPS public database.

2.3.1   Self-Assessment Manikin (SAM)

To assess the two dimensions of the affective space, the Self-Assessment Manikin (SAM), an affective rating system devised by Lang [27] was used to acquire the affective ratings. The SAM is a nonverbal pictorial assessment that is designed to assess the emotional dimensions (i.e. valence and arousal) directly by means of two sets of graphical manikins. The SAM has been extensively tested in conjunction with the IAPS and IADS and used in diverse theoretical studies and applications [3, 11, 15]. The SAM takes a very short time to complete (5 to 10 seconds). The SAM was reported to be capable of indexing cross-cultural results [28] and the results obtained using a Semantic Differential scale (the verbal scale provided in [29]). For using the SAM, there is little chance of confusion with terms as in verbal assessments. The SAM that we used was identical to the 9-point rating scale version of SAM that was used in [30], in which the SAM ranges from a smiling, happy figure to a frowning, unhappy figure when representing the affective valence dimension. On the other hand, for the arousal dimension, the SAM ranges from an excited, wide-eyed figure to a relaxed, sleepy figure.

Ratings are scored such that 8 represents a high rating on each dimension (i.e. positive valence, high arousal), and 0 represents a low rating on each dimension (i.e. negative valence, low arousal).
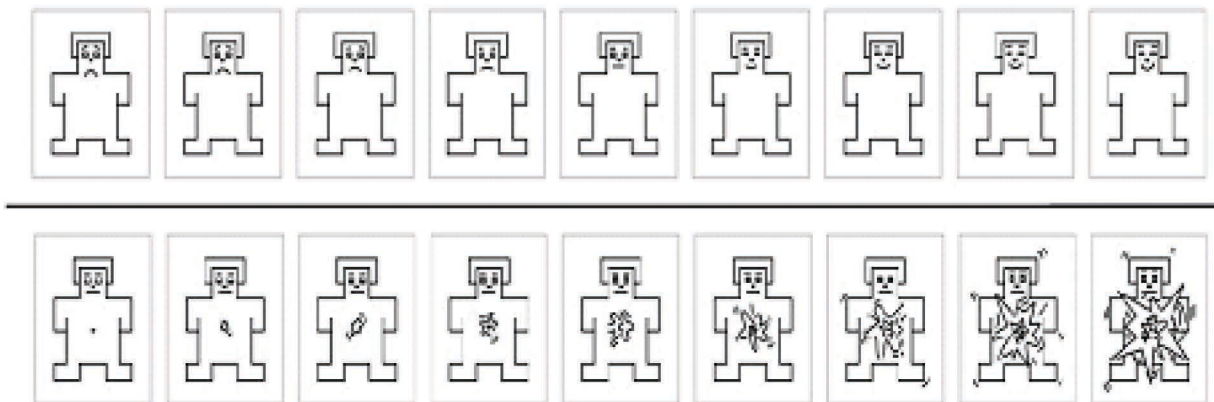
Figure 2. The SAM used in this study, in which the upper row represents valence and the lower row represents arousal

2.4 Extended Classifier Systems (XCSs)

2.4.1 Introduction

John Holland proposed Michigan-style Classifier Systems (CSs) in 1975, which is the prototype of the well-known classifier "LCS" [31]. LCS is a rule-based online learning algorithm, which incorporates Genetic Algorithm (GA) as a rule discovery component. Later on, Zeroth-level Classifier System (ZCS) was proposed to increase the understandability and performance [32]. ZCS adopts Q-learning (QL) like Reinforcement Learning (RL) component and retains the GA component. Finally, a classifier system known as XCS was proposed [8]. The XCS retain the QL and GA components in the ZCS but the fitness value of the rules in the XCS is referred to the accuracy of each rule on predicting payoff. Due to the stable performance and the capability to generalize extracted rules, XCS gained more attention from the main stream of research than other classifiers since it has been proposed.

On the research path of XCS, several versions of XCS have been developed to suit different needs required for real world applications. While the original XCS provides only binary string as input string (condition input) and with single discrete value as output, the XCS with real-value input (XCSR) which allows the XCS to accept continuous input have been proposed [33-35]. Lanzi suggested adding internal memory to XCS (named XCSM) for coping with complex non-Markovian environments [36], XCSI reduces the size of evolved classifier population [37], DXCS as a parallel version of XCS enhances the scalability of XCS [38]. To have XCS coping with function approximating tasks, Wilson himself also provided his idea of having continuous value as output in 2002 [39]. This idea is recently extended to a more advanced version of XCS named "Extended Classifier System for Function approximation task (XCSF)" [40].

Due to the development of the internet and WWW protocols, researchers nowadays are able to gather huge amount of data from the internet. To extract information from these data, data warehouse and data mining techniques have become popular research area. ML algorithms have been adopted widely for data mining tasks. XCS itself as one of the most important classifier has also been customized to fulfill the requirement of new tasks such as knowledge discovery and structure identification [41, 42] (e.g., probabilistic CS [43]). One of the most important feature that XCS provides to data mining tasks is its nature of being an online learning algorithm [44]. That is, XCS is able to adapt to dynamic environment, and even adapt to the environment in some extreme cases [45].

In practice, it is obvious that the characteristics of most of systems change gradually from time to time. The phenomena exist not only in the "seasonal changes" and unexpected structural changes in stock market, but also in the area of biomedical engineering (e.g. Heart Rate Variability (HRV) indexes computed from Electrocardiogram (ECG) that are usually used by psychophysiologists and the physicians in hospital for estimating subject's physiological state) and other fields. The characteristic of a system varies from time to time can make classification tasks more difficult for traditional su-
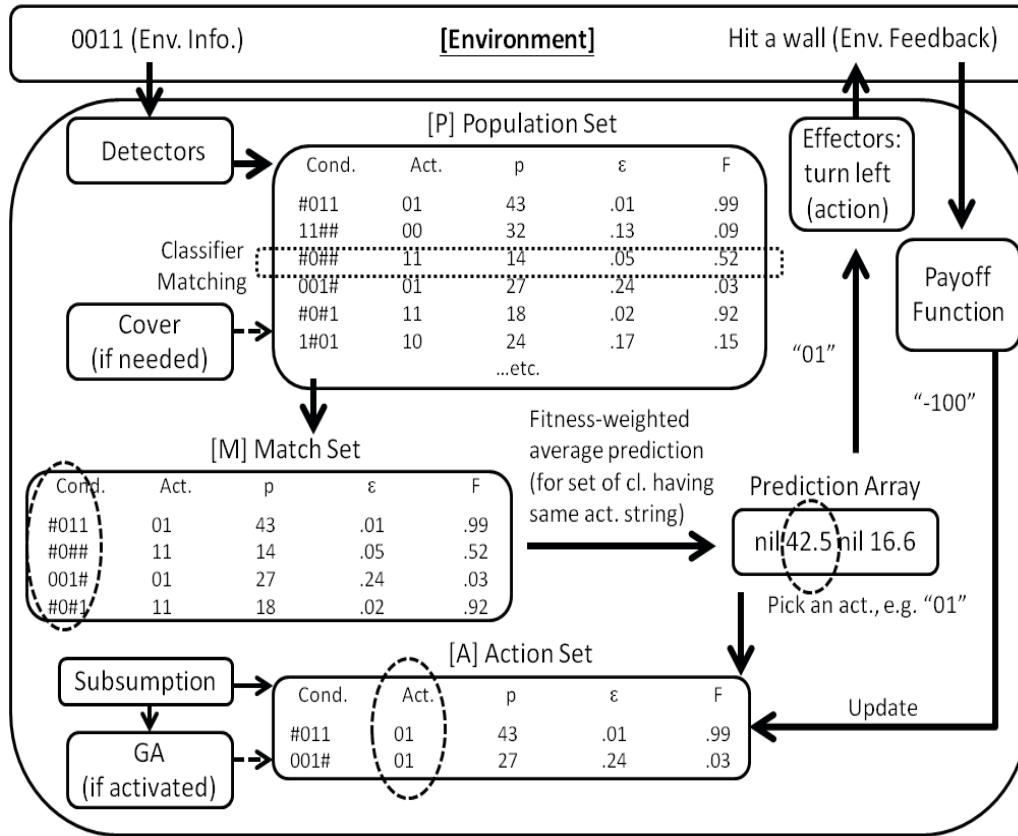
Figure 3. System architecture of XCS (for single-step problem)

pervised learning algorithm that does not provide online learning mechanism.

The XCS has been applied to wide range of classification tasks [46, 47] and is proved to be competitive for pattern recognition [6]. A considerable amount of literature describing applications based on XCS has been published in the area of security [48], finance [49, 50], medical research [51-53] and chip design [54, 55]. In the area of finance, XCS is known for its capability of financial time series forecasting [56-58]. The XCS was also used for developing personalized desktop by solving user-context classification tasks [43, 59].

2.4.2 XCS Classifier System

Although the XCS can be applied for both single-step problems and multi-step problems [8], this section focuses on describing only the mechanisms of XCS in solving single-step problems (i.e. classification task) instead of multi-step problems for simplicity. The flow of a typical XCS learning iteration is presented as follows: first, a detector obtains the environmental input (i.e. a binary string) at the beginning of a typical iteration, and uses the string for the matching process (see upper

left portion of Figure 3); second, during a classifier matching process, the XCS searches for classifiers in [P] which the covering condition space that represented by a condition string (0, 1, # for each bit, # indicates a bit that should be ignored, also termed "don't care" bit) includes the binary string input. All of the matched classifiers are placed into a match set (represents by [M]). If the [M] does not meet the predefined criterion [8, 60, 61] (usually related to the level of coverage of the suggested output (action string)), XCS applies a mechanism termed "cover" to generate new classifiers of which condition string matches the input binary string and action string is chosen at random; third, the XCS calculates the fitness weighted average prediction P_i from each set of classifiers that suggests a same output i (i.e. suggesting a same action) after [M] is generated; fourth, all P_i s are used to form a prediction array (PA) for output selection process. The action-selection regime is usually set to occasionally pick up an output i, which owns the maximal predicted payoff (i.e., max (P_i)) in the PA, and in the other time pick up an output randomly for exploration purpose; fifth, the XCS performs an action based on the selected output; and finally, after performing the action to the environment, a payoff function then

determines payoff (i.e. a numerical value) for XCS to update the classifiers. The payoff function is a function predefined by the user of XCS, to interpret the environmental feedback into numeric form of payoff (as a reward or punishment). The XCS uses the payoff in a RL component to update parameters (p, e, and F) of each classifier. The update process is held in only an action set (represents by [A]). The [A] is a set of classifiers (see bottom portion of Figure 3) that all suggest outputting i. The classifiers in [A] are the classifiers responsible for the payoff (determined from the environmental feedback) caused by the performed action. During the update process, the Rule (Classifier) Discovery part of XCS (see left bottom portion of Figure 3), that is, the GA is triggered occasionally to search for potentially accurate classifiers in the classifier (condition-action string representation) space. In addition, XCS performs subsumption in both the update process and GA, to enable "Macro-Classifier"'s (classifiers that are more general than others) to subsume other classifiers in order to reduce the number of redundant, overlapped classifiers. The remainder of this chapter provides details on the critical components of XCS.

2.4.3   RL Component

The RL component of XCS applies a QL-style update to the parameters of each classifier in [A] during the update process. First the XCS updates the prediction payoff p base on the received payoff using $p \leftarrow p + \beta(R - p)$, in which R represents the received payoff and β represents learning rate $(0 < \beta \leq 1)$. Second, the prediction error ε is updated using $\varepsilon \leftarrow \varepsilon + \beta(|R - p| - \varepsilon)$. Third, the XCS updates the fitness value F (used for the classifier space searching done by using the GA). The F in XCS is defined based on the accuracy of a classifier. Hence, in calculating F, the XCS calculates an accuracy value first using

$$\kappa = \begin{cases} 1, & \text{if } \varepsilon < \varepsilon_0 \\ \alpha\left(\frac{\varepsilon}{\varepsilon_0}\right)^{-\nu}, & \text{o.w.} \end{cases}$$

Equation 1 Calculation of accuracy value of classifiers in the XCS

in which the κ is set as 1 when ε is smaller than ε $(\varepsilon_0 > 0)$ to tolerate a classifier that contains prediction error if the prediction error is below ε_0. The value of κ decreases substantially (depending on the

settings of the parameter $\alpha\,(0 < \alpha < 1)$ and the exponent $\nu\,(\nu > 0))$ when a classifier's value of ε increases. After the update of κ, the XCS compares it's κ to that of other classifier's (i.e. other classifiers in the same [A]) by calculating its classifier's relative accuracy $\kappa' = \kappa / \sum_{x \in [A]} \kappa_x$. Finally, the F of the classifier is updated using $F \leftarrow F + \beta(\kappa' - F)$.

2.4.4   Rule Discovery Component Using GA

The discovery component in XCS searches for accurate classifiers in the classifier space by generating new classifiers using GA. The XCS applies the GA to an [A] when the average elapsed time since the last GA performed on all the classifiers in [A] is greater than q_{GA}. In GA, XCS selects two parent classifiers with a probability proportional to their fitness values. Two offspring classifiers are generated by applying crossover and mutation on the copies of their parents. Most parameters of the parents are inherited by their offspring, except for several parameters that must be initialized; for example, the fitness F, to be relatively pessimistic about the quality of the offspring, is multiplied by 0.1. After mutation, the generated offspring are inserted into the population.

2.4.5   Macroclassifiers

The XCS extracts generalized rules (classifiers) by reducing redundant classifiers. The idea of macroclassifier is implemented by using an additional parameter termed numerosity *num*. A classifier with numerosity *num* = n is equivalent to n regular classifiers. When XCS generates a new classifier, [P] is scanned to examine whether a macroclassifier exists with the same condition and action as that of the new classifier. If [P] has a classifier with the same condition and action, the value of *num* of the existing classifier (i.e. a macroclassifier) with the same condition and action is incremented by one instead of inserting the new classifier into [P]. Otherwise, the new classifier is added to the population with *num* set to one. Similarly, when a macroclassifier experiences a deletion, the value of *num* is decremented by one and the macroclassifier with numerosity *num* = 0 is removed from the [P]. The macroclassifier technique reduces redundant classifiers and also speeds up the XCS in generating [M].

2.4.6 Classifier Deletion and Subsumption

XCS removes classifiers from [P] if the sum of all *num* s of the classifiers in [P] exceeds a limit N (i.e. maximum population size predefined by the user of XCS) when inserting a new classifier into [P] (by either using the cover mechanism or GA). The probability of removing a classifier from [P] is proportional to the estimation of the size of [A]s (i.e. the [A] that the classifier usually appears in). The XCS also increases the probability of deletion of an experienced classifier with the value of F that is substantially lower than the average value of F of all classifiers in [P]. Subsumption deletion is a method to improve the generalization capability of XCS, and occurs after the update process of [A] and G. Hence, the subsumption is also called action set subsumption and GA subsumption [60]. During an action set subsumption, XCS selects an experienced classifier G with $\varepsilon < \varepsilon_0$ first; then G subsumes all the other classifiers in [A] that are less general than G, and the *num* of G is incremented based on the *num* s of the classifiers subsumed. The XCS also operates GA subsumption when new classifiers (i.e. offspring) are generated through the GA. The offspring are compared to their parent and subsumed if the parent classifier is experienced (defined by the times appeared in [A]) and more general.

3 HSV Patterns in the Affective Image Classification

3.1 Literature Review

To predict emotions of subject induced by an image, in 2005, Mikels et al. firstly categorized images in IAPS into different categories to identify images that are especially excellent in inducing emotions of subjects [62]. Later, a pioneering study [19] on affective image classification reported by Wu et al. applied the Support Vector Machine (SVM) on identifying the relationships between visual features extracted from images, and the semantic differential features (i.e. terms that given to subjects to describe the onset image, such as beautiful-ugly, dynamic-static, and tense-relaxed). The accuracy rate obtained in the study was relatively high (i.e. 80%); however, there was only one subject involved in the experiment, and the emotional state of the subject was implicitly estimated through seman-tic differential terms. To demonstrate the feasibility in affective image classification, subsequently, experiments with larger sample size (around 15 to 20 people) was conducted in [21, 23]; in these studies, emotions were explicitly defined as discrete emotional states, such as happy, surprising, sad, and angry. Further examinations on various features in affective image-classification task was reported by Machajdik in [22]; however, the obtained accuracy rates were relatively low (around 65%) in the between subject analysis in [21-23]. Latest findings in [11, 12] highlighted the drawback on using discrete emotion models, in which definition on emotions using "terms" may be vague and inaccurate for the subjects, and the use of discrete emotion model is generally application dependent, which may bias the collected dataset and the performance of classification model built. Furthermore, the use of discrete definitions also makes the experiment results hard to reproduce and hard to compare internationally. Hence, this study argues that the affective image classification studies should be conducted based on dimensional emotion model to reduce the difficulties in reproducing comparable results.

To clarify the objectives, the affective image classification problem is formatted into a system identification task (see Figure 4), the aim of the problem is to identify how human subjects interpret the affective characteristics of a given image; for example, to identify the human subject response by discovering rules, or training intelligent systems to predict the response (currently most of the works aimed on the later approach).

To evaluate the emotion elicitation of a subject, despite numerous approaches are available. For example, self-report [63], facial expression [64], keystroke dynamics, user data, and psychophysiological data [65]. This study decided to utilize self-report as the measurement tool, because self-report as a ground truth is considered to be more meaningful in the proposed problem and also the related future applications.

This study conducts an experimental study on affective image classification by adopting dimensional emotion model instead of applying discrete emotion models that were typically used in previous studies [22]. The SAM was used in this study to estimate the emotion elicitation of subjects in the perspectives of dimensional emotion
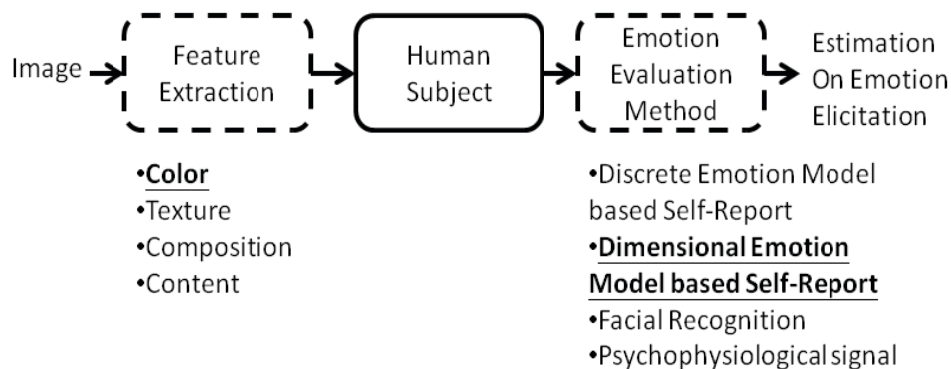
Figure 4. The affective image classification problem as a system identification task

model. The use of the experimental paradigm of dimensional emotion model, on the other hand, extends the traditional discrete emotion classification task into a continuous function approximation task, hence, in this study, the performance of classifiers were judged by Root Mean Square Error (RMSE) for two-dimension affective space prediction, and Mean Absolute Error (MAE) for one-dimension affective space prediction, instead of accuracy rate.

The objective of the study is to upgrade the performance of affective image classification. However, the results of previous studies, that obtained from experimental designs based on discrete emotional model, are hard to compare, and even incomparable in our case (because of the shift of performance criteria). Hence, this study examines the performance obtained from the described task, and hopes to provide a baseline for future study instead.

3.2 Experimental Setup

3.2.1 Subjects

There were 16 university subjects participated in the study (15 subjects is the typical sample size required in the field of affective image classification studies [19, 23]), ranging in age between 20 and 28 (M = 23.44, SD = 2.19; 10 men, 6 women). All subjects reported they were healthy, with no history of brain injury, cardiovascular problems, had normal or corrected-to-normal vision, and normal range of finger movement.

3.2.2 Experimental Procedure

To build an intelligent system that could predict the emotions of subjects elicit by image, a hu-

man subject experiment was conducted. The entire experiment conducted in this study complies the IAPS protocol of emotion inducement described in [24] to guarantee the effectiveness of the emotion induction procedure, and the clarity of the experimental design for reproduction. During the experiment, the subjects were requested to look at a screen which sequentially presents images and to correspondingly rate these images presented, by using computer-based SAM (through the use of mouse). The duration of the experiment was 10 minutes for each subject. Each trial (i.e. presentation of an image) started by presenting an image and displayed it for 6 seconds, then presented the SAM on the screen for the subject to manually rate the affective characteristics (i.e. self-report the induced emotion) of the presented image. The SAM was followed by a 15 s delay to ensure the emotional status of subject return to baseline before the start of next trial and a reasonable length to keep the subjects involved in the experiment.

3.2.3 Images Used

This study utilizes 20 images selected from IAPS [66] database in complying the IAPS image set selection protocol described in [66]. The image ids of the used images are as follows: 1120, 1310, 1390, 1710, 1720, 2160, 2220, 2520, 2530, 2540, 3160, 3220, 3250, 4300, 4460, 4470, 4660, 4750, 5950, 8160, 8200, and 9250. These images can be found in the IAPS database [66] using the ids listed above. The order of the image presentation was randomized to eliminate the effects due to the presentation sequence.

3.2.4 Environment Setting

The images were presented using a general PC with 32-inch (81.28 centimeters) monitor. The subjects were sat in a comfortable bed at a distance of approximately 1.5 meters away from the monitor in an EMI shielding room (Acoustic Inc. US) in which eliminates most of noise interferences and electrical noises. The CO_2 concentration of the environment was monitored during the entire experiment to guarantee reasonable CO_2 concentration (500 ppm ~ 1,300 ppm) to keep subjects sustain their attention during the experiment.

3.3   Method

3.3.1   HSV Model

This chapter adopts the approach of feature extraction similar to the former studies [22, 23], in which only basic features based on colors were extraction from the image (HSV model, in our case) instead of applying content based analysis, to eliminate the individual difference. Texture information was not used in this study because of the documentary-style natural of the IAPS images; images in the IAPS hold similar texture properties, and the related features extracted from IAPS images was reported useless in [22].

The HSV model is a cylindrical-coordinate representation commonly used in the area of computer graphics in replacement of RGB color model to obtain more intuitive values. In the HSV model, H represents hue, S represents saturation, and V represents value. Ordinarily, images stored in electronic devices such as personal computer are represented by a $M \times N$ matrix, in which the color of each element is displayed using RGB color model. The RGB model is a model consists of three coordinates as following: R represents red values, G represents green values and B represents blue values; red, green and blue are mixed together in a cube. For affective features analysis, features extracted from HSV model provide a more perceptually relevant representation on images.

– Hue

Hue is simply the attribute represents visual sensation on various colors similar to red, green, blue, or combinations of them. The value of hue is in the

interval of $0°$ and $360°$ (normalized to interval $[0, 1]$ in this study). The transformation from RGB to H is demonstrated as following: Firstly, normalizes R, G, and B of the target element into the interval $[0, 1]$. Secondly, calculates M, m and C from the normalized R, G, B.

$$M = \max(R, G, B)\,;\, m = \min(R, G, B)\,;\, C = M - m$$

Equation 2 The transformation from RGB to H (Step 2)

Thirdly, calculates H' and H.

$$H' = \begin{cases} 0, & \text{if } C = 0 \\ \frac{G-B}{C} \bmod 6, & \text{if } M = R \\ \frac{B-R}{C} + 2, & \text{if } M = G \\ \frac{R-G}{C} + 4, & \text{if } M = B \end{cases}$$

Equation 3 The transformation from RGB to H (Step 3)

$$H = 60° \times H'$$

Equation 4 The transformation from RGB to H (Step 4)

– Saturation

Saturation represents the level of colorfulness relative to its own brightness. The value of saturation is in the interval $[0, 1]$.

$$S = \begin{cases} 0, & \text{if } C = 0 \\ \frac{C}{M}, & \text{o.w.} \end{cases}$$

Equation 5 The calculation of saturation from the C and the M value

– Value (Brightness)

The Value (brightness) represents the brightness level relative to the brightness of a similarly illuminated white, defined as the largest component of the RGB color of an element (i.e, M, $0 \leq M \leq 1$) to form a hexagonal pyramid out of the RGB cube by projecting all three primary colors and the secondary colors such as cyan, yellow, and magenta into the new plane.

3.3.2   XCSF

The study applies XCSF to the affective image-classification task to cope with any possible non-linear characteristics contained in the target dataset. The XCSF is an extension of the XCS, a machine learning system based on Michigan-Style CSs. In 2002, XCSF, as a version of XCS used for function approximation was proposed [39]. The XCSF allows both real value inputs and real value outputs. In addition, the version of XCSF implemented in [67] allows multiple outputs The input accepts real value by using rotating hyperrectangle and rotating hyperellipsoid for condition representation [33, 68]. On the other hand, instead of selecting a discrete value as output according to fitness-weighted prediction value, the classifiers in the XCSF directly map the desire output using the prediction value produced by the linear approximation (i.e. $h\left(\vec{x}\right)=\vec{\omega}\ \vec{x}$ in which $\vec{x}$ represents the input vector and $\vec{\omega}$ represents weight vector). Each classifier in the XCSF updates its weight vector using Recursive Least Squares (RLS) method [68]. For performing the RLS, each classifier manage by XCSF updates its weight vector using

$$\vec{\omega} \leftarrow \vec{\omega}+\vec{k}\left[y_t-\left(\vec{x^*}-\vec{m^*}\right)^T\vec{\omega}\right]$$

Equation 6 Weight vector update of the XCSF based on the performance of RLS

where y_t represents target output, and $\vec{k}$ represents the gain vector computed by

$$\vec{k}=\frac{V^T\left(\vec{x^*}-\vec{m^*}\right)}{\lambda+\left(\vec{x^*}-\vec{m^*}\right)^T V^T\left(\vec{x^*}-\vec{m^*}\right)}$$

Equation 7 The calculation of the gain vector in the XCSF

The l (usually $0 \leq \lambda \leq 1$) used in Equation 6 and Equation 7 represents the forget rate of RLS. The lower the value of l is the higher the forget rate. The value of l is set to 1.0 for having an infinite memory (mostly used in time invariant problems). The matrix V hold by each classifier updates recursively using

$$V^T=\lambda^{-1}\left[I-\vec{k}\left(\vec{x^*}-\vec{m^*}\right)^T\right]V^T$$

Equation 8 The update of matrix V in the XCSF

The fitness value used for the GA in the XCSF is the relative classifier accuracy calculates from system error [60]. For further detail, sufficient information about XCS can be found in Butz's algorithmic description of XCS [60], and also the recent advances in XCSF [33, 39, 68]. To summarize, the XCSF can be understood as a manager which manages a set of classifiers. Each of the classifiers maps from a subspace in the feature space to the landscape-function output using a linear-fitting method.

3.3.3   Model Building

The workflow of the preprocessing and model building are provided in Figure 5 The workflow of the built prediction model ; the preprocessing of the image data was based on HSV model without applying content based analysis, 6 features were used for model building in this study, including: average hue, standard deviation of hue, average saturation, standard deviation of saturation, average brightness, and standard deviation of brightness. The model was built to predict the induced emotion rated by subjects in terms of valence and arousal through SAM. The prediction of valence and arousal can be real number herein according to the definition of valence and arousal in the dimensional theory of emotion [11]. To avoid overfitting problem, a Leave-One-Out-Cross-Validation (LOOCV) on leaving one sample at each time for testing set and the remain samples for training set, which is the standard practice for analyzing limited dataset, was used for building the model.

For details on the setting of LR and XCSF for building the models, the LR analysis was done by using the Weka implementation of data mining tools [69], in which Akaike criterion was used for model selection and M5's method was used for attribute selection; all the co-linear attributes were excluded.

The XCSF used in this study was adopted from the Java implementation version on XCSF contributed by Stalph and Butz (2009) [67]. For parameters setting, a = 1.0; b = 0.1; d = 0.1; l = 1.0; q_{GA} = 50; e_0 = 0.5; d_{rls} = 1000; q_{del} = 20; c = 1.0; m = 1.0; q_{sub} = 20; the GA subsumption was turned on. Although the maximal population size N was set to 6,400~10,000 to maximize the performance of XCSF, the number of classifiers quickly converged
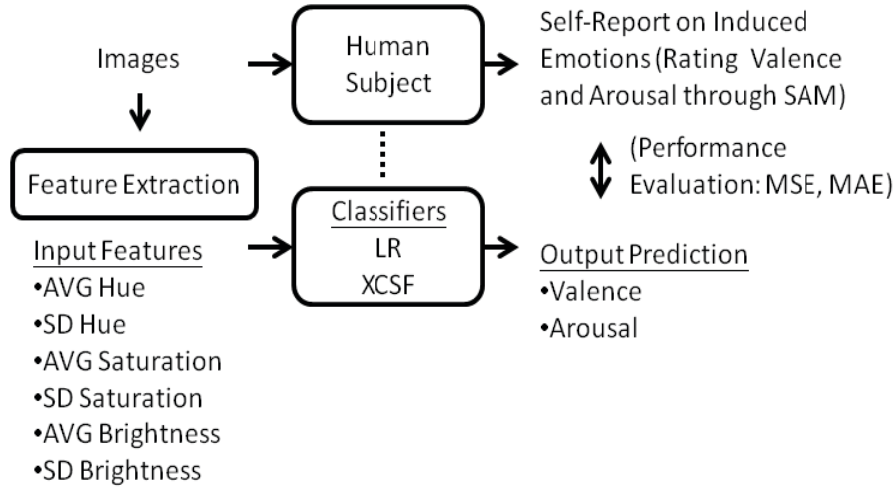
Figure 5. The workflow of the built prediction model

to 5,400 during the model training. To examine the performance of the system, e_0 was set to various values. However, it appears relatively small effect on the learning performance in regard to the learning speed and system error. During the model training, the XCSF was sequentially presented with 20,000 instances randomly selected from the training dataset.

3.4 Results and Discussion

3.4.1 Collected Dataset

The collected dataset contains 20 images (1024x768 JPEG) used in the experiment, and the image affective ratings rated by 16 subjects through SAM. The experiment totally acquired 320 rows (actually, 318 rows, while two rows were excluded due to machine mal functioning) of raw data (images, and the affective ratings of the images, 20 rows for each subject). Figure 6 presents the distribution of the ratings selected by subjects on all images; it is observed that most subjects were aroused with either unpleasant feelings or pleasant feelings by the displayed images, no obvious skewed was observed in the distribution of valence (histogram was examined but not shown).

3.4.2 Model Performance and Discussion

The performance in regard to RMSE/MAE and the standard deviation of MAEs (represents by SD) achieved by LR and XCSF are provided in.

Table 1. The performance achieved by distinct classifiers

Prediction Results	Affective Dimension				
	Valence		Arousal		(Valence, Arousal)
Method	MAE	SD	MAE	SD	RMSE
uniRand	2.569	(N/A)	2.530	(N/A)	(N/A)
largCount	1.613	(N/A)	1.480	(N/A)	(N/A)
LR	1.482	1.021	1.481	1.070	2.564
XCSF	0.970	0.747	1.460	1.029	2.165

The manner of calculating RMSE and MAE are provided as following:

$$\text{RMSE} = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\left[(V_i - VP_i)^2 + (A_i - AP_i)^2\right]}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|V_i - VP_i| \text{ or } \frac{1}{N}\sum_{i=1}^{N}|A_i - AP_i|$$

Equation 9 The calculation of RMSE and MAE in this research

in which N represents sample size; V_i and A_i represents the values of the valence and arousal corresponds to the i-th sample; and VP_i and AP_i represents the system prediction on the values of valence and arousal corresponds to the i-th sample. The MAEs are used here to evaluate the performance of a built model in predicting valence and arousal. The RMSE is always adopted when a classifier is used for predicting valence and arousal in pairs, and the MAEs are always adopted when a classifier is used for predicting valence and arousal separately.
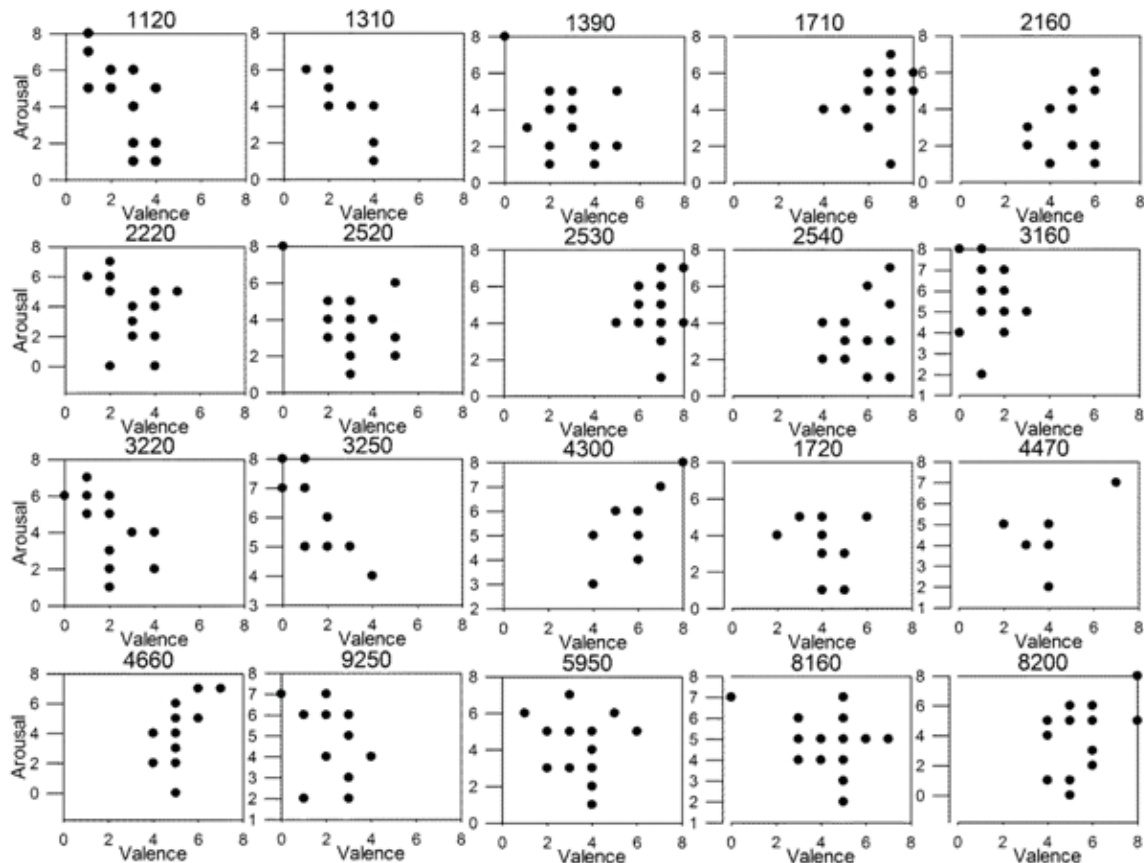
Figure 6. The distribution of the induced emotion of subjects on all images

While the emotional ratings are not uniformly distributed, the MAE of prediction can be artificially underestimated; hence, two models, 1) uniRand: making predictions in a uniformly random manner, and 2) largCount: making constant predictions based on the weighted-average valence, and weighted-average arousal, based on the ratings, in which average value of valance was nearly 3.931 and average value of arousal was nearly 4.349, were introduced to compare the MAE achieved by LR and XCSF. The performance of distinct classifiers is provided in Table 1. The performance of LR on predicting valence in regard to MAE is 1.483, which is relatively low while the uniRand achieved only 2.570 and largCount achieved 1.614. In addition, the MAE value achieved by XCSF on predicting valence decreased the MAE value achieved by LR from 1.483 ± 1.02 to 0.971 ± 0.747 (-0.512), demonstrates the capability of XCSF on mapping functions that possibly contain non-linearity by managing a set of linear classifiers. The MAE achieved by XCSF was small and the standard deviation of the MAE is tolerable.

To further examine the performance of XCSF

on this task, the performance of classifiers on predicting valence values in terms of MAE are illustrated in Figure 7, in which x-axis represents valence and y-axis represents MAE. The MAE on each valence is represented by four bars: the MAE achieved by uniRand, largCount, XCSF, and LR, respectively. The MAE achieved by XCSF is smaller than the MAE of LR, uniRand and largCount at most of the emotional ratings; the MAE of XCSF is only larger than uniRand at the rating with the largest count. A skew on MAE was observed for those ratings that represent for "being pleasant", that is, 5~8, possibly due to the sample size of the rating, while in Table 1, the numbers of samples of valence equals to 5, 6, 7 are larger than the numbers of samples of valence equals to 0 and 1.

Conversely, the MAE of XCSF in valence 0 and 8 are also high. The finding suggests that insufficient on sample size of a class may lead to low performance of XCSF on approximating the corresponding output value even the training instances were selected from the training dataset randomly during the XCSF iterative training process. However, the MAE of XCSF at valence value equals to

4.0 is not the lowest, indicating that the larger sample size may only guarantee the efficacy of XCSF in function approximation but not eliminating all existing errors. We observed that approximately 30 samples (which is nearly 10% of the number of raw data), is sufficient for XCSF to approximate a valence value in the collected dataset. Otherwise, for example, the MAEs of XCSF on predicting the samples that valence equals to 0 and 8 are relatively high, in which the numbers of samples were small. To explain the extreme cases happened in valence 0 and valence 8 in the psychological perspective, we've noticed that some of the subjects reported that they tend to select ratings in the middle of the scale rather than selecting those ratings represent for extreme cases. Such phenomenon may cause the non-linearity of the distances between the levels of valence and arousal. Further clarification is required for this issue; a well-designed transformation may be adequate.
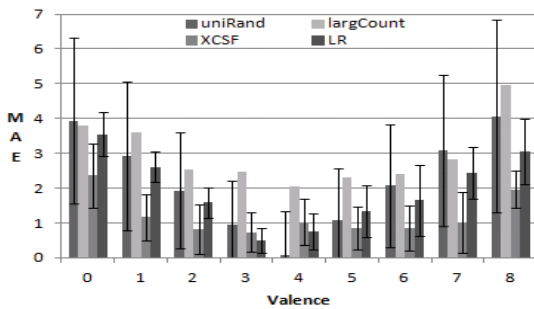

Figure 7. The MAE of XCSF on each valence value. The standard deviation of MAE made by XCSF is nearly 0.54~0.91.
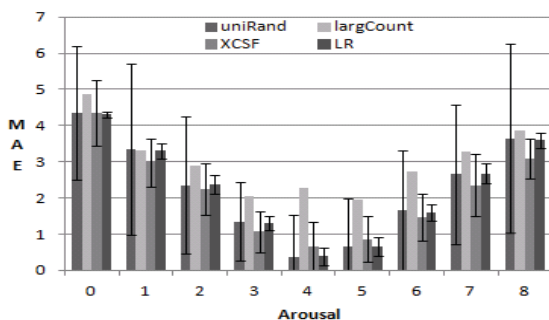

Figure 8. The MAE of XCSF on each arousal value. The standard deviation of MAE made by XCSF is nearly 0.27~0.91

Figure 8 also presents similar phenomenon. The MAEs achieved by XCSF in each level of arousal mostly outperform the MAEs of LR. However, in general, the MAEs achieved by XCSF at each level were increased, and the decreases in error are not sufficiently significant. The results provide that MAE of LR and XCSF did not outperform largCount in the prediction on arousal. Such observation indicates that the prediction model of arousal built by LR and XCSF did not adequately identify the problem structure, possibly due to the ineffectiveness of SAM on estimating subjects' arousal, while some of subjects reported that during the experiment the definition of "being aroused" can be easily confused with the definition of the tendency of valence. The confusion was possibly caused by the cultural difference, but similar results were not highlighted previously in the main stream of the research.

To further identify the discovered knowledge, the prediction models built by LR are provided in Equation 10 and Equation 11.

Valence = -2.3147 * Avg_Saturation + 4.6681 * Avg_Brightness + 12.8186 * SD_Brightness - 0.3798 Equation 10 The model of predicting valence based on HSV properties Arousal = 3.4657 * SD_Saturation - 3.3625 * SD_Brightness + 4.3804

Equation 11 The model of predicting arousal based on HSV properties

From Equation 10, the saturation of an image tends to lower the valence, whereas the brightness tends to enhance the pleasant feelings, and the effect of standard deviation of brightness on valence is even more substantial. On the other hand, the affective characteristics of an image making people feel aroused, is negative correlated with the standard deviation of brightness, whereas the standard deviation of saturation increases the effect.

Gender information was not used for model building because the effect of gender is eliminated by the use of IAPS protocol, the use of gender information did not substantially improve the performance during this study as well. By contrast to the gender information, within subject analysis was also applied to the dataset using LR and XCSF. The result indicates that without using the content within the image, the effects of individual difference is relatively small (not shown here).

In regard to the concerns of the application of this study, the RGB model is device-dependent, due

to the color elements (such as phosphors or dyes) and their response to the individual R, G, and B levels vary from manufacturer to manufacturer, different devices may detect or reproduce a given RGB value distinctively, or even in the same device over time. Such characteristics may cause variations in the proposed experimental results. However, similar problem may also occur in other studies that utilize IAPS, but currently, of our best knowledge, no previous study reports further problem due to it.

4 Spatial-Frequency Patterns in the Affective Image Classification

4.1 Literature Review

To index the affective characteristics of images, an intuitive approach is to have a large number of people manually rate all the images and calculate descriptive statistics from the ratings. On the other hand, recent studies utilize color, texture, and composition information of images; also the application of content analysis, to achieve affective image classification [22, 70].

However, the features related to spatial-frequency domain that are proven to be useful for pattern recognition have not been explored yet. In addition, contributed by recent advances in methodology, the resolution in frequency analysis has been improved. Hence, this chapter aims to solve the affective image-classification task by using the features related to spatial-frequency domain, and the XCSF [68] (i.e. one of a latest version of LCS [71]). The dataset used for the classification task is collected from a human-subject experiment conducted in our laboratory. The performance of the built intelligent machine in performing affective image-classification task was validated by 10-Fold CV. The proposed method may be applied to other images in real-word.

4.2 Experimental Setup

4.2.1 Subjects

There were 16 university subjects participated in the study (15 subjects is the typical sample size required in the field of affective image classification studies [19, 23]), ranging in age between 20 and 28 (M = 23.44, SD = 2.19; 10 men, 6 women). All sub-

jects reported they were healthy, with no history of brain injury, cardiovascular problems, had normal or corrected-to-normal vision, and normal range of finger movement.

4.2.2 Experimental Procedure

To build an intelligent system that could predict the emotions of subjects elicit by image, a human subject experiment was conducted. The entire experiment conducted in this study complies the IAPS protocol of emotion inducement described in [24] to guarantee the effectiveness of the emotion induction procedure, and the clarity of the experimental design for reproduction. During the experiment, the subjects were requested to look at a screen which sequentially presents images and to correspondingly rate these images presented, by using computer-based SAM (through the use of mouse). The duration of the experiment was 10 minutes for each subject. Each trial (i.e. presentation of an image) started by presenting an image and displayed it for 6 seconds, then presented the SAM on the screen for the subject to manually rate the affective characteristics (i.e. self-report the induced emotion) of the presented image. The SAM was followed by a 15 s delay to ensure the emotional status of subject return to baseline before the start of next trial and a reasonable length to keep the subjects involved in the experiment.

4.2.3 Images Used

This study utilizes 20 images selected from IAPS [66] database in complying the IAPS image set selection protocol described in [66]. The image ids of the used images are as follows: 1120, 1310, 1390, 1710, 1720, 2160, 2220, 2520, 2530, 2540, 3160, 3220, 3250, 4300, 4460, 4470, 4660, 4750, 5950, 8160, 8200, and 9250. These images can be found in the IAPS database [66] using the ids listed above. The order of the image presentation was randomized to eliminate the effects due to the presentation sequence.

4.2.4 Environment Setting

The images were presented using a general PC with 32-inch (81.28 centimeters) monitor. The subjects were sat in a comfortable bed at a distance of approximately 1.5 meters away from the monitor in an EMI shielding room (Acoustic Inc. US) in which

eliminates most of noise interferences and electrical noises. The CO2 concentration of the environment was monitored during the entire experiment to guarantee reasonable CO2 concentration (500 ppm ˜ 1,300 ppm) to keep subjects sustain their attention during the experiment.

4.3 Method

4.3.1 Two-Dimensional Hilbert-Huang Transform (2D-HHT)

Spatial-frequency analysis on images is one of the well-known techniques used in the field of image processing and computer vision [72, 73]. The information in frequency domain was found abundant by physiologists [74]. It was found that various spatial-frequencies can lead to distinct characteristics of visual stimulations. Moreover, the orientation of visual stimulation can cause different efficacies in stimulations of cortical receptors [75, 76].

Traditionally, Fast Fourier Transform (FFT) is used to transform an image into frequency domain. However, due to the assumption of that series of target data should be at least piecewise stationary, the FFT-based techniques (e.g., spectrogram), is not suitable for modeling local phenomena or when higher resolution is required. Hence, recently HHT was proposed to obtain higher frequency resolution toward Instantaneous Frequency (IF) [77]. Later, the use of such concept in spatial-frequency analysis was also reported [78]. The HHT is a two-phase transformation, which firstly apply an Empirical Mode Decomposition (EMD) on the target data series to extract Intrinsic Mode Functions (IMFs). Secondly, Hilbert Transform (HT) is applied to each IMF to obtain required frequency domain information (i.e., IF). The EMD is a shifting process that can be used to extract IMFs from a data series $X(s)$. The IMF is defined as a monocomponent by satisfying the criterias as following:

1 has the number of zero crossings and extrema one difference at most,

2 symmetric with respect to the local mean, and

3 the $X(s)$ should has at least two extrema.

After the procedure of EMD, n IMFs, namely, IMF1, IMF2, IMF3, ..., IMFn, and the

residuals(r_n), denoted as

$$X(s) = \sum_{j=1}^{n} C_j + r_n$$

Equation 12 The decomposition of an input signal based on the EMD provided in Equation 12, are extracted from $X(s)$. The residuals (r_n) is a data series which is the remainder series of target data series after the EMD shifting process removes all the IMFs from the original target data series.

The procedure of EMD, different from the Fourier and Wavelet Decomposition, is fully data-driven. By being adaptive and unsupervised, the EMD improves the efficiency of signal decomposition and can be applied to the non-linear and non-stationary signal (details on the procedure of the EMD please refer to [77]). After the EMD, the HT is then applied to each IMF

$$Y_j(s) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{C_j(\tau)}{s-\tau} d\tau$$

Equation 13 Transforming each IMF$_j$ extracted by the EMD into $Y_j(s)$

Each IMFj can be represented by the conjugate pair of $Y_j(s)$ and $C_j(s)$, hence can be represented by an analytical signal $Z(s) = C_j(s) + iY_j(s) = a(s)e^{iq(s)}$, in which the amplitude

$$a_j(s) = \sqrt{C_j(s)^2 + Y_j(s)^2}$$

Equation 14 The computation of amplitudes in HT and phase $q_j(s) = \arctan(Y_j(s)/C_j(s))$. Based on the definition stated above, the IF$_j$ can be derived by applying a derivative on $q_j(s)$ (i.e. $w_j = dq_j(s)/ds$). Then, an analytical representation of $X(s)$, can be derived

$$X(s) = \sum_{j=1}^{n} a_j(s) \left[i \int \omega_j(s) ds \right]$$

Equation 15 Analytical representation of the input signal based on the HHT

Originally, the EMD was proposed to decompose one-dimensional data. To construct a 2D-HHT, the concept of EMD was extended to 2D in this study based on the concept listed as follows:

1 identify the extrema (maxima and minima) of the image by sliding a 3-by-3 grid;

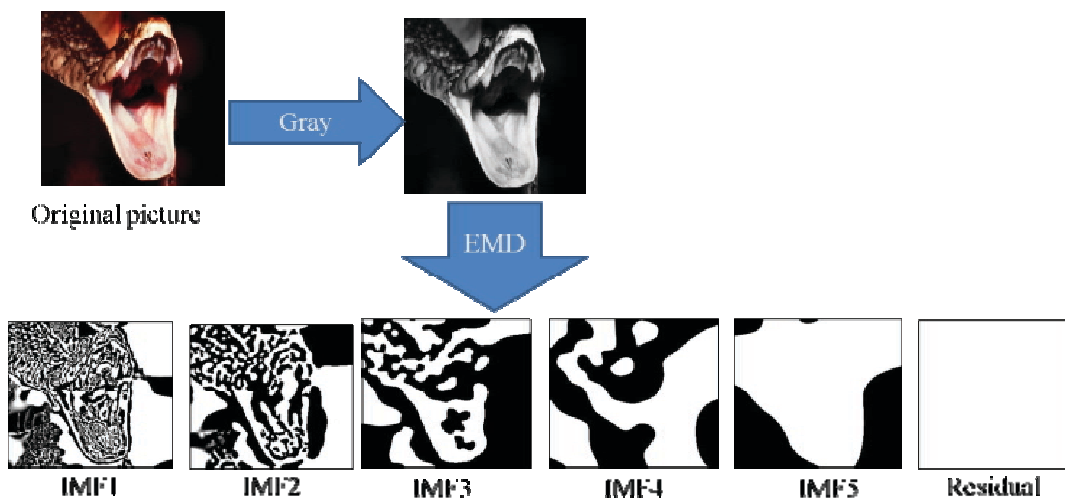2 generate two smooth 2D surfaces to fit the found maxima and minima;

Figure 9. The illustration of the data processing in this study, the application of 2D EMD on IAPS picture 1120

3 compute the local mean by averaging two surfaces; and

4 the equation of applying 2D-EMD then can be rewrite from Equation 12, to

$$f(x, y) = \sum_{j=1}^{n} C_j(x, y) + r_n(x, y).$$

Equation 16 The decomposition of a 2D input signal based on the 2D-EMD

The data processing done in this study is illustrated in Figure 9. The original image (1024x768 resolutions) was first down-sampled to 128x128 resolutions. The color setting was changed from RGB color into gray color. Second, the 2D EMD is applied to the image. For extracting IFs from IMFs, this study applies the concept of partial HT by applying 1D HT to each orientation (i.e. each row and each column) and unit, in order to extract spatial-frequency features that account for different orientations of visual stimulations [75]. The IF analysis method used in this study was inspired by the work in [79] which provides a show case on estimating the changes of IF data series. This study mainly adopts three indexes as follows: 1) $F\_Q\_IMF_j$ represents frequency value in the 1st quarter of the histogram area of IF_x; 2) $A\_I\_IMF_j$ represents the ratio between the 1st and the 2nd halves of the histogram area of IF_j; 3) $M\_I\_IMF_j$ represents the ratio between the maxima found in the 1st and 2nd halves of the histogram area of IF_j. This study applies totally 12 features listed below as follows: 1) The vertical side (the direction of applying 1D

HT) $F\_Q\_IMF_1$, the horizontal side (the direction of applying 1D HT) $F\_Q\_IMF_1$, the vertical side $A\_I\_IMF_1$, the horizontal side $A\_I\_IMF_1$, the vertical side $M\_I\_IMF_1$, the horizontal side $M\_I\_IMF_1$; 2) the vertical side $F\_Q\_IMF_2$, the horizontal side $F\_Q\_IMF_2$, the vertical side $A\_I\_IMF_2$, the horizontal side $A\_I\_IMF_2$, the vertical side $M\_I\_IMF_2$, the horizontal side $M\_I\_IMF_2$.

We totally acquired 318 rows of the feature vector from the collected data set. The method that we use to build the prediction model is introduced in the following section.

4.3.2 Model Building

Models were built to predict the emotion ratings rated by subjects in terms of valence and arousal through SAM. The prediction of valence and arousal can be real number herein according to the definition of valence and arousal in the dimensional theory of emotion [11]. Besides the XCSF, this study also applies several well-known machine-learning techniques for comparison purpose. Zero-R is a majority voting learning scheme that predicts the majority class in any data set. In a classification task, the Zero-R classifies an instance into the majority class, whereas in a prediction task, the Zero-R predicts the mean value of all the instances. Thus, the performance of the Zero-R can be considered as a baseline performance of the classification class, which should be beaten by any algorithm that learns decision boundaries from the data set without over-fitting. One-layer method such as LR [80] and

multi-layered method with transfer function such as Radial-Basis-Function (RBF) Network [69] were used in this study. The LOOCV which leaves one sample out at each time as a testing set and the remaining samples as a training set was used for model building.

The XCSF used in this study was adopted from the Java implementation version on XCSF contributed by Stalph and Butz (2009) [67]. For parameters setting, a = 1.0; b = 0.1; d = 0.1; l = 1.0; q_{GA} = 50; e_0 = 0.5; d_{rls} = 1000; q_{del} = 20; c = 1.0; m = 1.0; q_{sub} = 20; the GA subsumption was turned on. Although the maximal population size N was set to 6,400~10,000 to maximize the performance of XCSF, the number of classifiers quickly converged to 5,400 during the model training. To examine the performance of the system, e_0 was set to various values. However, it appears relatively small effect on the learning performance in regard to the learning speed and system error. During the model training, the XCSF was sequentially presented with 20,000 instances randomly selected from the training dataset.

4.4   Results and Discussion

4.4.1   Collected Dataset

The collected dataset contains 20 images (1024x768 JPEG) used in the experiment, and the image affective ratings rated by 16 subjects through SAM. The experiment totally acquired 320 rows (actually, 318 rows, while two rows were excluded due to machine mal functioning) of raw data (images, and the affective ratings of the images, 20 rows for each subject). Figure 6 presents the distribution of the ratings selected by subjects on all images; it is observed that most subjects were aroused with either unpleasant feelings or pleasant feelings by the displayed images, no obvious skewed was observed in the distribution of valence (histogram was examined but not shown).

4.4.2   Model Performance and Discussion

The performance evaluation based on MAE and the standard deviation (represents by SD) of the MAEs achieved by the methods used is provided in Table 2. The ZeroR represents the Zero-R classifier, LinearReg represents the LR model, and RBFNet represents the RBF network. The number of nodes

(clusters) of the RBF network was set to 200 based on the result of the examination on the performance changes caused by the number of nodes.

The MAE of a prediction model which predicts at random on the value of valence and arousal is 4.0. Hence, the MAE 1.453±1.076 achieved by the LR seems to be fair.

Table 2. The performance achieved by benchmark classifiers and XCSF

Prediction Results		Affective Dimension	
		Valence	Arousal
Method	Statistics		
ZeroR	MAE	1.617	1.491
	SD of MAE	1.110	1.065
LinearReg	MAE	1.453	1.427
	SD of MAE	1.076	1.083
RBFNet	MAE	0.950	1.471
	SD of MAE	0.747	1.021
XCSF	MAE	0.950	1.461
	SD of MAE	0.755	1.011

The MAE achieved by the RBF network is 0.949±0.747, which further shows a reduction of the error by 35%. This result indicates the existence of the non-linearity characteristic of the dataset collected. The MAE achieved by the XCSF was 0.950±0.755. The equivalence in the performance of RBF network and XCSF indicates the capability of XCSF on mapping non-linear functions. The mechanism of the XCSF in model building by managing a set of linear classifiers seems to be comparable to the multi-layered based method with non-linear transfer function. To further examine the performance of the XCSF, the MAEs of the XCSF on each valence and arousal value are also provided in Figure 10 and Figure 11. To compare the MAE achieved by the XCSF, the MAEs that achieved by uniRand, a classifier that makes predictions in a uniformly random manner are also included in these figures.

The performance of the XCSF in predicting each valence value is illustrated in Figure 10 in which x-axis represents the valence value, y-axis represents the MAE value. The MAEs that the classifiers achieved on each valence are represented by three bars. The right most bar represents the MAE achieved by the XCSF. The MAEs of uniRand and

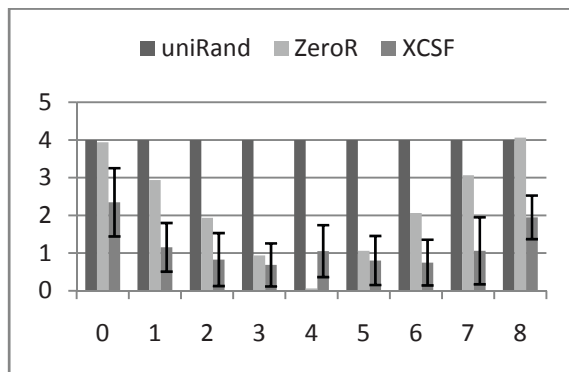ZeroR are represented by the first and the second bar.

Figure 10. The MAE of XCSF on each valence value. The standard deviation of MAE made by XCSF is nearly 0.58~0.91. Based on the SAM ratings, the maximal MAE is 8 and minimal MAE is 0.
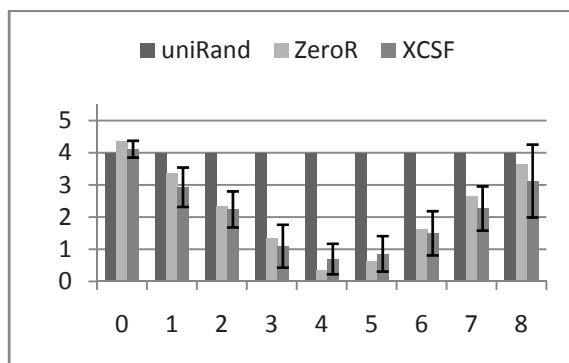

Figure 11. The MAE of XCSF on each arousal value. The standard deviation of MAE made by XCSF is nearly 0.26~0.69. Based on the SAM ratings, the maximal MAE is 8 and minimal MAE is 0.

The MAE achieved by the XCSF is smaller than the MAE achieved by uniRand and ZeroR at most ratings (i.e. the value of valence and arousal). The MAE of XCSF is only larger than the MAE of ZeroR at the ratings near the mean values. A skew on the value of MAEs is observed for the largest and lowest valence values (i.e. 0~1 and 8). This is possibly due to the sample size of these ratings, since the numbers of samples of valence equals to 0, 1, and 8 are smaller. This finding suggests that insufficient sample size of a class (e.g., valence = 8) may lead to bad performance of XCSF in predicting the corresponding output value. However, the MAE of XCSF at valence 4.0 was not the lowest, which indicates that the larger sample size only guarantees

the efficacy of XCSF in function approximation instead of eliminating all exist errors. In our observation, approximately 30 samples (which is nearly 10% of the number in our collected data set) is sufficient for the XCSF to build a model to predict a valence value in our collected dataset. On the other hand, this phenomenon happened in valence 0 and 8 could also be explained by a psychological approach. That is, some of the subjects reported that they tended to rate the values in the middle of the scale rather than those values that represent extreme emotional experiences. This may cause the non-linear characteristics of the distances between the levels of valence and arousal. Further clarification is required for this issue; an appropriate transformation may be applied to the data to improve the result.

Similar results can be found in Figure 11, in which for the ZeroR, the prediction was set to 3.937. The MAE achieved by the XCSF in each level of arousal substantially outperformed the MAE of uniRand and ZeroR. However, the Figure 11 shows the increase of the MAEs achieved by the XCSF at each level of arousal. This could be explained by that most subjects reported that during the experiment, they confused the definition of "being aroused" with "the tendency of valence". The reaction of the subjects is possibly caused by the cultural difference, but similar results were not highlighted previously in the research community that applies the IAPS and SAM.

Gender information was not used for model building, but the within-subject analysis was conducted. However, we found that without the applying content analysis to the image, the effect of individual difference is relatively small. To further examine the performance of the XCSF in this task, the ROC curve of the
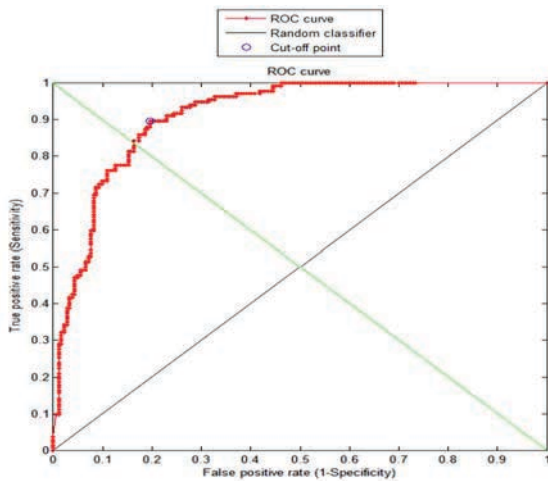
Figure 12. The ROC curve made by XCSF on prediction of the event when valence of the image is rated smaller than 4 (neutral).
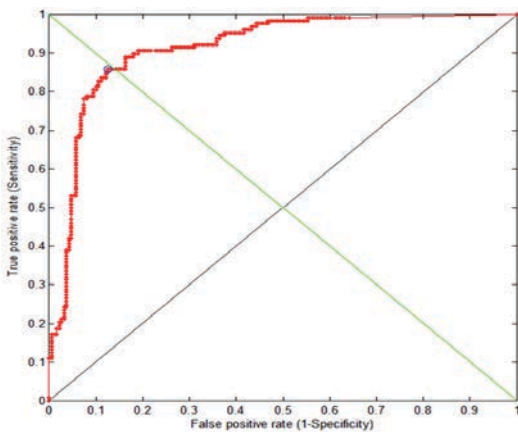

Figure 13. The ROC curve made by XCSF on prediction of the event when valence of the image is rated larger than 4 (neutral)

XCSF in predicting the value of valence of an image being rated smaller or larger than 4 (i.e. 4 is the value of valence that represents a neutral state) are provided in Figure 12 and Figure 13. The result shows that the AUC achieved by the XCSF is significantly (p-value < .000) larger than the random classifier the AUC of predicting valence < 4 is 0.913, in which the AUC of predicting valence > 4 is 0.914). However, the AUC achieved by the XCSF in predicting the value of arousal being smaller or larger than 4 is relative small (p-value < .005). The AUC of predicting arousal < 4 is 0.594 and the AUC of predicting arousal > 4 is 0.573. The accuracy rates achieved by the XCSF on the cut-off point are: 84.3% (for predicting valence < 4), 86.8% (valence > 4), and 53.1% (arousal < 4), and 56.0%

(arousal > 4). In addition, the ROC curve of RBF network was also examined because the MAE made by RBF network was favorable in comparison with the MAE made by XCSF. The results are provided in Table 3.

Table 3. The results of ROC Curve Produced by XCSF and RBF network

Prediction Target Method Estimation		V< 4	V> 4	A< 4	A> 4
XCSF	AUC	0.913	0.914	0.594	0.573
	Accuracy	84.30%	86.80%	53.10%	56.00%
RBFnet	AUC	0.682	0.673	0.495	0.505
	Accuracy	70.50%	65.20%	50.10%	63.30%

V: Valence, A: Arousal.

To further identify the extracted knowledge, the prediction models built by the LR are provided in Equation 17 and Equation 18.

Valence = 3.2893 * F_Q_IMF1_col + 0.3651 * F_Q_IMF1_row + 2.6606 * A_I_IMF1_row + 0.4394 * F_Q_IMF2_row - 0.2629 * A_I_IMF2_col + 2.335 * A_I_IMF2_row- 1.0522

Equation 17 The model of predicting valence based on spatial-frequency properties

Arousal = -2.4297 * F_Q_IMF1_col - 0.9253 * F_Q_IMF1_row + 0.1896 * A_I_IMF1_col - 1.2495 * A_I_IMF1_row + 0.4578 * M_I_IMF1_col - 0.3721 * A_I_IMF2_col + 0.1047 * M_I_IMF2_col + 7.0513

Equation 18 The model of predicting arousal based on spatial-frequency properties

For building models by the LR, Akaike criterion was used for model selection and the M5's method was used for attribute selection, in which all the co-linear attributes were excluded. The equations show that F_Q_IMF1_col, F_Q_IMF1_row, and A_I_IMF2_row were the main factors that affect the affective ratings during the experiment. The F_Q_IMF1_col, F_Q_IMF1_row, and A_I_IMF2_row show positive relationship to the rating of valence. These results indicate that the stimulations from horizontal side are more effective than the stimulations from vertical side. The horizontal side of the image may contain abundant information. Conversely, the affective characteristics of an image in

regard to making people feel aroused, is negative correlated with F_Q_IMF1_col and A_I_IMF1_row. In addition, the offset of the equation 3 is +7.0513. These results indicate that the effects of activation in motivational system due to a visual stimulus are influenced by the asymmetric of cortical receptors responsible for distinctive directions of the spatial-frequency visual stimulations.

5   Conclusion

The overall goal of this study was to build an intelligent machine that can classify images based on their affective characteristics, especially the classification based on features extracted from the spatial-frequency domain. To achieve this goal, two novel affect detectors, two speed-up techniques for the XCS, and a novel 2D approach of applying the HHT method were developed. The developed systems were built and validated using multiple human-subject experiments and compared with the existing related systems. The rest of this chapter presents the achieved objectives, main conclusions from each contribution chapter, and the future work that stems from this research work.

5.1   Achieved Objectives

The following research objectives have been fulfilled by this work to achieve the overall research goal.

– For the first time, high resolution spatial-frequency features were extracted from the given images and used in building an affective classification model. By utilizing the proposed novel 2D feature-extraction method, the developed algorithm readily demonstrated spatial-frequency calculation at a resolution that existing 2D-FFT, wavelet-based methods, and HHT cannot.

– For the first time, conducted controlled experiments on this issue that adopt standard instruments which make the results cross-cultural and comparable to future studies which follow the same standard. The use of the dimensional theory of emotions in this study enabled the use of rich standard methodology and paradigms for conducting experiments. When without the use

of these methods, the obtained results were hard to compare and reproduce, leading the demonstrated techniques controversial.

In addition to achieving the above established research objectives, this work provided a detailed investigation and analysis of the models built for classifying images. This analysis revealed that no matter which the images are and who the subjects are, the strength of a specific frequency band in a specific directions on the influence of the affective characteristics could be the same (i.e. image independent and user independent). Further, the introduced 2D-HHT method are not simply another way of obtaining spatial-frequency features as the proposed methods fundamentally change the way that an ordinary 2D-FFT functions can do for improving resolution (e.g., windowing). Also that a standard 2D-FFT does not use all available amplitude information and do not decompose the given image to extract frequency domain information, whereas the developed HHT based 2D method effectively exploit the hidden information in an image by using EMD.

5.2   Main Conclusions

This section presents the main conclusions and highlights from the two major contribution chapters (Chapter 3 and Chapter 4).

On the patterns in the affective image classification, two models were built and validated based on multiple human-subject experiments. These models were built based on HSV properties of images and the features extracted from spatial-frequency domain through a 2D HHT method. By using the proposed 2D HHT method, this study obtains high resolution information in the spatial-frequency domain. The result indicates that both the models show comparable results to the results that were reported in the previous studies.

The XCSF was demonstrated to adequately approximate the output landscape of the affective image-classification problem in the collected datasets. The relationships between the used features and the affective characteristics of images were examined and shown in Equation 10, Equation 11, Equation 17, and Equation 18. The property of images in the HSV and the spatial-frequency domain is proved to be influential to the affective

characteristics of images. These user-independent results are favorable because less object-detection technique is required for model building. Moreover, there should be less possible interference due to individual difference which should have been eliminated during the model-building process.

5.3 Future Work

For suggested future work, applications on indexing affective characteristics of the images on the internet, or providing feedbacks to the users to improve the quality of life; for example, to ease people, or to excite people; are possible future work. In future, the results produced by this study should be replicated using images obtain from sources other than the IAPS database, to validate the experimental results in regard to generality. Further examination on the mechanisms and pathway between the affective information contained in the spatial-frequency domain and the cortical receptors in human eye, is also suggested.

5.4 Closing Remarks

This research work has shown that images can be indexed based on their affective characteristics. The 2D nature and the complexity in the images added additional difficulties to the task. The use of HSV models and high-resolution 2D spatial-frequency feature-extraction method such as the 2D-HHT can led the systems to build accurate, maximally general and compact models in indexing various IAPS images as well as pictures in real-world. Effectively exploiting the combined power of 2D-HHT and the XCSF, various real-value modeling and function approximation problems could be solved in a simple and straight forward manner. Understanding of the affective image-classification task reveals that the standard spatial-frequency feature extraction methods do not exploit all available information embedded in the amplitude matrix of images, whereas the developed 2D-HHT based systems effectively exploit the embedded-features and spatial information during the model-building process. This study has shown a new perspective of indexing images based on their affective characteristics, not just a new feature extraction method for indexing images, which is needed to replicate human capabilities and should lead to various novel applications.

6 Acknowledgement

References

[1] Kuo, W.J., et al., Intuition and Deliberation: Two Systems for Strategizing in the Brain. Science, 2009. 324(5926): p. 519-522.

[2] Chowdhury, R.M.M.I., G.D. Olsen, and J.W. Pracejus, Affective Responses to Images In Print Advertising: Affect Integration in a Simultaneous Presentation Context. Journal of Advertising, 2008. 37(3): p. 7-18.

[3] Chang, C., The Impacts of Emotion Elicited By Print Political Advertising on Candidate Evaluation. Media Psychology, 2001. 3(2): p. 91-118.

[4] Kyung-Sun, K., Effects of emotion control and task on Web searching behavior. Information Processing & Management, 2008. 44(1): p. 373-385.

[5] Mitchell, T.M., Machine learning. 1997. Burr Ridge, IL: McGraw Hill, 1997. 45.

[6] Orriols-Puig, A., J. Casillas, and E. Bernad-Mansilla, Genetic-based machine learning systems are competitive for pattern recognition. Evolutionary Intelligence, 2008. 1(3): p. 209-232.

[7] Russell, S. and P. Norvig, Artificial Intelligence: A Modern Approach (2nd Edition)2002: Prentice Hall.

[8] Wilson, S.W., Classifier fitness based on accuracy. Evol. Comput., 1995. 3(2): p. 149-175.

[9] Picard, R.W., Affective Computing2000, Cambridge MA: The MIT Press.

[10] Ortony, A. and T. Turner, What's basic about basic emotions. Psychological review, 1990.

[11] Bradley, M.M., Emotional memory: a dimensional analysis, in Emotions: Essays on emotion theory, S.H.M.v. Goozen, N.E.v.d. Poll, and J.A. Sergeant, Editors. 1994, Lawrence Erlbaum: Hillsdale, NJ. p. 97-134.

[12] Bradley, M.M. and P.J. Lang, Emotion and motivation, in Handbook of Psychophysiology, J.T. Cacioppo, L.G. Tassinary, and G. Berntson, Editors. 2007, Cambridge University Press: New York, NY. p. 581-607.

[13] Lang, P.J., The motivational organization of emotion: Affect-reflex connections, in Emotions: Essays on emotion theory1994, Lawrence Erlbaum: Hillsdale, NJ. p. 61-93.

[14] Lang, P.J., The Emotion Probe - Studies of Motivation and Attention. American Psychologist, 1995. 50(5): p. 372-385.

[15] Bolls, P.D., A. Lang, and R.F. Potter, The Effects of Message Valence and Listener Arousal on Attention, Memory, and Facial Muscular Responses to Radio Advertisements. Communication Research, 2001. 28: p. 627-651.

[16] Antonio R, D., Emotion in the perspective of an integrated nervous system. Brain Research Reviews, 1998. 26(2–3): p. 83-86.

[17] Bechara, A., The role of emotion in decision-making: Evidence from neurological patients with orbitofrontal damage. Brain and cognition, 2004. 55(1): p. 30-40.

[18] LaBar, K.S. and R. Cabeza, Cognitive neuroscience of emotional memory. Nat Rev Neurosci, 2006. 7(1): p. 54-64.

[19] Wu, Q., C. Zhou, and C. Wang, Content-Based Affective Image Classification and Retrieval Using Support Vector Machines, in Affective Computing and Intelligent Interaction, J. Tao, T. Tan, and R. Picard, Editors. 2005, Springer Berlin / Heidelberg. p. 239-247.

[20] Joshi, D., et al., Aesthetics and Emotions in Images. Signal Processing Magazine, IEEE, 2011. 28(5): p. 94-115.

[21] Liu, N., et al., Associating Textual Features with Visual Ones to Improve Affective Image Classification, in Affective Computing and Intelligent Interaction, S. D'Mello, et al., Editors. 2011, Springer Berlin / Heidelberg. p. 195-204.

[22] Machajdik, J. and A. Hanbury, Affective image classification using features inspired by psychology and art theory, in Proceedings of the international conference on Multimedia2010, ACM: Firenze, Italy. p. 83-92.

[23] Zhang, H., et al., Analyzing Emotional Semantics of Abstract Art Using Low-Level Image Features, in Advances in Intelligent Data Analysis X, J. Gama, E. Bradley, and J. Hollmn, Editors. 2011, Springer Berlin / Heidelberg. p. 413-423.

[24] Lang, P.J., M.M. Bradley, and B.N. Cuthbert, International affective picture system (IAPS): Affective ratings of pictures and instruction manual, 2008: University of Florida, Gainesville, FL.

[25] Sanchez-Navarro, J., et al., Psychophysiological, behavioral, and cognitive indices of the emotional response: A factor-analytic study. Spanish Journal of Psychology, 2008. 11(1): p. 16-25.

[26] Kensinger, E.A., R.J. Garoff-Eaton, and D.L. Schacter, Effects of emotion on memory specificity: Memory trade-offs elicited by negative visually arousing stimuli. Journal of Memory and Language, 2007. 56(4): p. 575-591.

[27] Lang, P.J., Behavioral treatment and bio-behavioral assessment: Computer applications, in Technology in Mental Health Care Delivery Systems, J. Sidowski, J. Johnson, and T. Williams, Editors. 1980, Ablex Pub. Corp.: Norwood, NJ. p. 119-137.

[28] Morris, J.D., Observations: SAM: the Self-Assessment Manikin; an efficient cross-cultural measurement of emotional response. Journal of advertising research, 1995. 35(6): p. 63-68.

[29] Mehrabian, A. and J.A. Russell, An approach to environmental psychology1974, Cambridge, MA: the MIT Press.

[30] Bradley, M.M. and P.J. Lang, The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective ratings of sounds and instruction manual. University of Florida, Gainesville, FL, Tech. Rep. B-3, 2007.

[31] Holland, J.H., Adaptation in Natural and Artificial System1992, Cambridge, MA, USA: MIT Press.

[32] Wilson, S.W., ZCS: A Zeroth Level Classifier System. Evolutionary Computation, 1994. 2(1): p. 1-18.

[33] Wilson, S.W., Get Real! XCS with Continuous-Valued Inputs. Learning Classifier Systems, 2000. 1813: p. 209-219.

[34] Stone, C. and L. Bull, For Real! XCS with Continuous-Valued Inputs. Evolutionary Computation, 2003. 11(3): p. 299-336.

[35] Dam, H.H., H.A. Abbass, and C. Lokan, Be real! XCS with continuous-valued inputs, in Proceedings of the 2005 workshops on Genetic and evolutionary computation2005, ACM: Washington, D.C. p. 85-87.

[36] Lanzi, P.L. Adding memory to XCS. in IEEE World Congress on Computational Intelligence. 1998.

[37] Wilson, S.W., Compact Rulesets from XCSI, in Advances in Learning Classifier Systems, P. Lanzi, W. Stolzmann, and S. Wilson, Editors. 2002, Springer Berlin / Heidelberg. p. 65-92.

[38] Dam, H.H., H.A. Abbass, and C. Lokan, DXCS: an XCS system for distributed data mining, in Proceedings of the 2005 conference on Genetic and evolutionary computation2005, ACM: Washington DC, USA. p. 1883-1890.

[39] Wilson, S.W., Classifiers that approximate functions. Natural Computing, 2002. 1(2): p. 211-234.

[40] Lanzi, P.L., et al., Generalization in the XCSF Classifier System: Analysis, Improvement, and Extension. Evol. Comput., 2007. 15(2): p. 133-168.

[41] Bull, L., E. Bernad-Mansilla, and J. Holmes, Learning Classifier Systems in Data Mining: An Introduction, in Learning Classifier Systems in Data Mining, L. Bull, E. Bernad-Mansilla, and J. Holmes, Editors. 2008, Springer Berlin / Heidelberg. p. 1-15.

[42] Butz, M., et al., Knowledge Extraction and Problem Structure Identification in XCS, in Parallel Problem Solving from Nature - PPSN VIII, X. Yao, et al., Editors. 2004, Springer Berlin / Heidelberg. p. 1051-1060.

[43] Muruzbal, J., A probabilistic classifier system and its application in data mining. Evol. Comput., 2006. 14(2): p. 183-221.

[44] Orriols-Puig, A., J. Casillas, and E. Bernad-Mansilla, First approach toward on-line evolution of association rules with learning classifier systems, in Proceedings of the 2008 GECCO conference companion on Genetic and evolutionary computation2008, ACM: Atlanta, GA, USA. p. 2031-2038.

[45] Dam, H., C. Lokan, and H. Abbass, Evolutionary Online Data Mining: An Investigation in a Dynamic Environment, in Evolutionary Computation in Dynamic and Uncertain Environments, S. Yang, Y.-S. Ong, and Y. Jin, Editors. 2007, Springer Berlin / Heidelberg. p. 153-178.

[46] Quirin, A., et al. Analysis and evaluation of learning classifier systems applied to hyperspectral image classification. in Intelligent Systems Design and Applications, 2005. ISDA '05. Proceedings. 5th International Conference on. 2005.

[47] Butz, M., et al., Effective and Reliable Online Classification Combining XCS with EDA Mechanisms, in Scalable Optimization via Probabilistic Modeling, M. Pelikan, K. Sastry, and E. CantPaz, Editors. 2006, Springer Berlin / Heidelberg. p. 249-273.

[48] Akbar, M.A. and M. Farooq, Application of evolutionary algorithms in detection of SIP based flooding attacks, in Proceedings of the 11th Annual conference on Genetic and evolutionary computation2009, ACM: Montreal, Canada. p. 1419-1426.

[49] Armano, G., A. Murru, and F. Roli, Stock Market Prediction by a Mixture of Genetic-Neural Experts. International Journal of Pattern Recognition and Artificial Intelligence (IJPRAI), 2002. 16(5): p. 501-526.

[50] Tsai, W.-C. and A.-P. Chen. Global Asset Allocation Using XCS Experts in Country-Specific ETFs. in Convergence and Hybrid Information Technology, 2008. ICCIT '08. Third International Conference on. 2008.

[51] Sprogar, M., M. Sprogar, and M. Colnaric, Autonomous evolutionary algorithm in medical data analysis. Computer Methods and Programs in Biomedicine, 2005. 80(Supplement 1): p. S29-S38.

[52] Passaro, A., F. Baronti, and V. Maggini, Exploring relationships between genotype and oral cancer development through XCS, in Proceedings of the 2005 workshops on Genetic and evolutionary computation2005, ACM: Washington, D.C. p. 147-151.

[53] Baronti, F., et al., Machine learning contribution to solve prognostic medical, in Outcome prediction in cancer, A. Taktak and A.C. Fisher, Editors. 2007, Elsevier.

[54] Bernauer, A., et al., Combining Software and Hardware LCS for Lightweight On-chip Learning, in Organic Computing — A Paradigm Shift for Complex Systems, C. Mller-Schloer, H. Schmeck, and T. Ungerer, Editors. 2011, Springer Basel. p. 253-265.

[55] Bernauer, A., et al., Autonomous multi-processor-SoC optimization with distributed learning classifier systems XCS, in Proceedings of the 8th ACM international conference on Autonomic computing2011, ACM: Karlsruhe, Germany. p. 213-216.

[56] Armano, G., NXCS Experts for Financial Time Series Forecasting, in Applications of Learning Classifier Systems, L. Bull, Editor 2004, Springer. p. 68-91.

[57] Chen, A.-P. and Y.-H. Chang. Using extended classifier system to forecast S&P futures based on contrary sentiment indicators. in Evolutionary Computation, 2005. The 2005 IEEE Congress on. 2005.

[58] Chen, A.-P., et al., Applying the Extended Classifier System to Trade Interest Rate Futures Based on Technical Analysis, in Proceedings of the 2008 Eighth International Conference on Intelligent Systems Design and Applications - Volume 032008, IEEE Computer Society. p. 598-603.

[59] Shankar, A. and S.J. Louis, XCS for Personalizing Desktop Interfaces. Evolutionary Computation, IEEE Transactions on, 2010. 14(4): p. 547-560.

[60] Butz, M. and S. Wilson, An Algorithmic Description of XCS, in Advances in Learning Classifier Systems, P. Luca Lanzi, W. Stolzmann, and S. Wilson, Editors. 2001, Springer Berlin / Heidelberg. p. 267-274. Wilson, S.W., Generalization in the XCS Classifier System, 1998.

[61] Mikels, J., et al., Emotional category data on images from the international affective picture system. Behavior Research Methods, 2005. 37(4): p. 626-630.

[62] Bradley, M.M. and P.J. Lang, Measuring emotion: the self-assessment manikin and the semantic differential. Journal of Behavior Therapy and Experimental Psychiatry, 1994. 25: p. 49-59.

[63] Cohen, I., et al., Facial expression recognition from video sequences: temporal and static modeling. Computer Vision and Image Understanding, 2003. 91(1–2): p. 160-187.

[64] Kim, K.H., S.W. Bang, and S.R. Kim, Emotion recognition system using short-term monitoring of physiological signals. Medical & Biological Engineering & Computing, 2004. 42(3): p. 419-427.

[65] Lang, P.J., M.M. Bradley, and B.N. Cuthbert, International Affective Picture System (IAPS), in Techinal Manual and Affective Ratings 1999, The Center for Research in Psychophysiology, University of Florida: Gainesville, FL.

[66] Stalph, P.O. and M.V. Butz, Documentation of JavaXCSF, 2009: Retrieved from University of Wurzburg, Cognitive Bodyspaces: Learning and Behavior website.

[67] Butz, M.V., P.L. Lanzi, and S.W. Wilson, Function Approximation With XCS: Hyperellipsoidal Conditions, Recursive Least Squares, and Compaction. Evolutionary Computation, IEEE Transactions on, 2008. 12(3): p. 355-376.

[68] Hall, M., et al., The WEKA data mining software: an update. SIGKDD Explor. Newsl., 2009. 11(1): p. 10-18.

[69] Lee, P.-M., Y. Teng, and T.-C. Hsiao. XCSF for prediction on emotion induced by image based on dimensional theory of emotion. in Proceedings of the fourteenth international conference on Genetic and evolutionary computation conference companion. 2012. ACM.

[70] Holland, J.H. and J.S. Reitman, Cognitive systems based on adaptive algorithms. SIGART Bull., 1977(63): p. 49-49.

[71] Li, S. and B. Yang, Multifocus image fusion using region segmentation and spatial frequency. Image and Vision Computing, 2008. 26(7): p. 971-979.

[72] Leonard, H., et al., Brief Report: Developing Spatial Frequency Biases for Face Recognition in Autism and Williams Syndrome. Journal of Autism and Developmental Disorders, 2011. 41(7): p. 968-973.

[73] John G, D., Two-dimensional spectral analysis of cortical receptive field profiles. Vision Research, 1980. 20(10): p. 847-856.

[74] Webster, M.A. and R.L. De Valois, Relationship between spatial-frequency and orientation tuning of striate-cortex cells. J. Opt. Soc. Am. A, 1985. 2(7): p. 1124-1132.

[75] Kobayashi, K., et al., Head and body sway in response to vertical visual stimulation. Acta Otolaryngologica, 2005. 125(8): p. 858-862.

[76] Huang, N., et al., The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. Roy. Soc. Lond. A, 1998. 454: p. 903-995.

[77] Tay, P.C. AM-FM Image Analysis Using the Hilbert Huang Transform. in Image Analysis and Interpretation, 2008. SSIAI 2008. IEEE Southwest Symposium on. 2008.

[78] Caseiro, P., R. Fonseca-Pinto, and A. Andrade, Screening of obstructive sleep apnea using Hilbert–Huang decomposition of oronasal airway pressure recordings. Medical Engineering & Physics, 2010. 32(6): p. 561-568.

[79] Kaw, A. and E. Kalu, Numerical Methods with Applications 2010: autarkaw.