

On the Dynamic Pattern Analysis, Discovery and Recognition

Ke Chen
School of Informatics
The University of Manchester
Manchester M60 1QD, United Kingdom
Email: Ke.Chen@manchester.ac.uk

1. Background

In the real world, most of objects can be treated as one kind of patterns in general. Although patterns are ubiquitous and of miscellaneous forms, they can be generally divided into two categories: static and dynamic patterns. In nature, static patterns tend to unchange dramatically upon generation whilst dynamic patterns always tend to change in a wide range. For instance, printed characters can be regarded as a class of static patterns since upon generation they are not altered considerably despite the slight deformation by the noise corruption. In contrast, handwriting artifacts of a word may vary significantly. On the one hand, different people have different writing styles, which results in different shapes for the same word. On the other hand, one could write the same word in various forms at different times. Thus, handwriting artifacts become a class of dynamic patterns. In terms of pattern analysis, discovery and recognition, dynamic patterns lead to more challenging problems than static patterns. In artificial intelligence, there are many challenging problems, ranging from biometrics to bioinformatics, relevant to various dynamic patterns, and their solutions rely on effective dynamic pattern analysis, discovery and recognition techniques [1].

As a prominent characteristic, the variability of dynamic patterns in the same class becomes gradually larger, which leads to tremendous difficulties in discriminating between dynamic patterns belonging to different classes. In addition, dynamic patterns often convey mixing information, which is often hardly separable, so that the direct use of the mixing information can result in the inadequate performance for a specific task due to the interference of irrelevant information. For example, it is well known that a speech signal, a type of dynamic patterns, generally contains three types of information; i.e., linguistic, speaker-specific, and environmental information [2]. A speech recognition task demands only linguistic information rather than speaker-specific and environmental information; an ideal system should recognize speech regardless of speakers. Conversely, a speaker recognition task needs speaker-specific information; an ideal speaker recognition system should carry out the voice-based authentication regardless of linguistic information or whatever a speaker utters. As a consequence, the direct use of all mixing information in either speech or speaker recognition hinders either of two systems from producing higher performance.

The characteristics of dynamic patterns give rise to several challenging problems as follows: how to generate a parsimonious yet robust representation of dynamic patterns that tolerates their intra- and inter-class variability, how to highlight the relevant information and simultaneously suppress affection of any irrelevant information, how to make use of intrinsic contextual information for dynamic pattern analysis and discovery, and how to create an appropriate learning model for dynamic pattern recognition to yield the consistently good generalization without being affected by the constant change of dynamic patterns?

For dynamic pattern analysis, discovery and recognition, the author and his collaborators have been systematically researching into this theme for over one decade. This article tends to give a brief overview of our previous endeavors in tackling dynamic pattern related problems and further presents our ongoing research topics.

2. Overview of Our Research

For over one decade, we have been systematically researching into dynamic pattern analysis, discovery and recognition. Our research can be summarized as three aspects; i.e., *exploration/exploitation of effective representations*, *use of the divide-and-conquer principle* and *exploitation of intrinsic contextual information*. The techniques developed have been applied to several real problems ranging from speaker recognition to computer vision.

2.1. On Representations of Dynamic Patterns

In pattern recognition, one of the most important topics is feature extraction that distills silent features from raw data in order to facilitate the next-stage processing and overcome the “curse of dimensionality” problem [1]. For dynamic patterns, feature extraction becomes critical; a *representation* or a feature set extracted from raw data needs to capture intra- and inter-class variability as well as highlight relevant features (suppress irrelevant features reciprocally) for a specific task.

Typically, speech is a type of dynamic patterns. In our early study [3], we investigated a perceptually-processing based speech representation for speaker recognition by introducing characteristics of the human auditory system for bearing variability of speech and a new transformation for highlighting the speaker-specific features. The proposed representation is based on the perceptual linear predictive analysis and an adaptive component weighting idea. A comparative study showed that our speech representation outperforms most of classical ones in the robustness against noise and highlighting speaker-specific information [3].

Although the best representation is always anticipated, it is often unavailable due to the nature of dynamic patterns. Instead, there are multiple representations extracted by different feature extraction methods that characterize dynamic patterns from various perspectives. Although each of different representations can be used individually, none of them is perfect to represent dynamic patterns completely. Our research uncovers that for a set of dynamic patterns, a single representation often characterizes only a subset of dynamic patterns very well but fails to present others [4]. It suggests that the simultaneous use of different representations provides an alternative way for better characterizing dynamic patterns.

For the simultaneous use of different representations, our initial endeavor simply constructed a composite representation by lumping different representations together [5]-[7]. Although the use of such a composite representation has yielded relatively better performance, it results in three problems as follows: 1) curse of dimensionality, 2) difficulty in formation, and 3) redundancy. These problems not only hinder the use of a composite representation from a significant improvement but also incur much higher computational cost (for an example in speaker recognition, see [4]).

In order to tackle problems mentioned above, our next endeavor was combining multiple classifiers trained on different representations [8], [9]. The basic idea is a two-stage learning procedure; multiple classifiers are first trained independently on different individual representations and then combined on the decision level to reach a final consensus. In our previous studies [8], we systematically investigated the state-of-the-art combination schemes, including the associative switch, the Bayesian fusion and the Dempster-Shafer fusion, on homogeneous and heterogeneous classifiers trained on different representations. Furthermore, we proposed a novel optimal linear scheme especially for combining multiple probabilistic classifiers on different representations [9]. In general, the use of different representations in this way yields the considerably better performance than either the use of an individual representation or the use of a composite representation [8], [9]. Unfortunately, there is also a weakness behind this methodology; i.e., different representations are used in an indirect way and the two-stage learning procedure results in only a sub-optimal solution that could be

sensitive to intra- and inter-class variability of dynamic patterns. In addition, the use of heterogeneous classifiers on different representations leads to a new problem, yet another combination or correspondence problem; that is, how to select a proper one from a set of candidate representations for a specific classifier so that the heterogeneous classifiers used for combination can always associate with their proper individual representations for the maximal synergy [8]. To our knowledge, this problem is still open up to now in general.

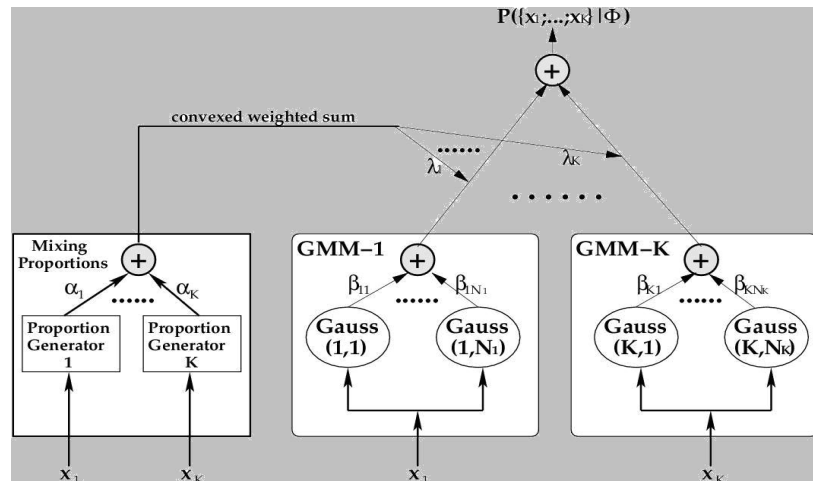


Fig. 1. Generalized Gaussian mixture model [4]: an implementation of the soft-competition scheme on different representations of dynamic patterns.

In order to better characterize dynamic patterns, we proposed a soft-competition scheme on different representations [4], [10], [11]. This scheme tends to overcome all the problems appearing in two aforementioned methodologies on the simultaneous use of different representations. The core idea behind this scheme is the use of a soft competition principle; i.e., given a dynamic pattern, its available representations compete each other for the right of representation and, unlike the winner-take-all principle, the soft-competition principle allows all the representations to work together so that for the current represented pattern, a winning representation would simply play a more important role. It is worth stating that the soft-competition of different representations takes place for every dynamic pattern and therefore a winning representation for the current pattern may be a loser for other patterns. In reality, it is impossible to know which representation is a winner in advance for a given pattern. Thanks to available training data, we cast the soft-competition problem as a learning task and develop a unified finite mixture model for soft-competition on different representations in a probabilistic sense; the unconditional probabilistic model for pattern modeling [4] and the conditional probabilistic model for classification [10], [11]. Thus, the learning task becomes a maximal likelihood problem for parameter estimation on training data. As illustrated in Fig. 1, a generalized Gaussian mixture model is designed to implement the unconditional finite mixture model. In this implementation, multiple Gaussian mixture models are employed to model dynamic patterns based on their different representations, whereas weights of different representations for a dynamic pattern are produced by a linear combination of the proportion generators working on different representation of the current dynamic pattern. This architecture forms an input-dependent model, which significantly distinguishes from those systems of combining multiple classifiers/models on different representations in an input-independent way. Another model was also designed for implementing the conditional finite mixture model on different representations [10], [11]. Both speaker modeling and recognition tasks demonstrate the usefulness of our soft-competition scheme [4], [10], [11].

In summary, our research mentioned above indicates that along with the exploration of a single representation for dynamic patterns for a specific task, the simultaneous use of

different representations, in particular, the optimal use of different representations like our soft-competition scheme, provides a generic and novel way to characterize dynamic patterns and often leads to the better performance.

2.2. On Use of the Divide-and-Conquer Principle

As a general principle, the divide-and-conquer strategy has been widely applied in various scientific fields to tackle a difficult yet complicated problem. As a consequence, an original problem is first divided into several smaller but simpler sub-problems that are subsequently solved independently. Afterwards, solutions to sub-problems are seamlessly integrated to form a complete solution. In our research, we have endeavored to apply the divide-and-conquer principle to dynamic pattern recognition from different perspectives.

Our initial endeavor was applying the mixture of experts [12] and its variants [13], [14] in speaker recognition. This model carries out the divide-and-conquer principle via its modular neural network architecture in a probabilistic sense. To a great extent, our applications of the mixture of experts [5] and its variants [15] manifest that such models is capable of dealing with the intra- and inter-class variability [5], [15]. In order to improve the performance of original model in dynamic pattern recognition, we have developed several variants, including the introduction of time-delay [6], [7] and input-weighted schemes [16] to the original model, and the development of alternative learning algorithms [17]. Our variants and new learning algorithms yield the better performance in dynamic pattern recognition, as exemplified by speaker identification [6], [7], [16], [17] and other benchmark tasks [17].

For use of the divide-and-conquer principle in recognition, a critical problem is how to divide a problem into a number of appropriate sub-problems corresponding to available learners, which can be view as model selection from another perspective; i.e. how to find the right number of local learning models for a given problem. In this context, we have come up with several novel approaches. First of all, we adopt the boosting methodology [18] to solve the model selection problem by developing a novel boosting algorithm that allows component learners to have different architectures [19]. As an application of our boosting algorithm, simple input/output HMMs [14] of different topological structures are employed to deal with different type of variability in a sequential way so that their ensemble can tackle a recognition problem without the need of explicit model selection. Simulations have demonstrated the effectiveness of our approach in sequence classification [19].

For supervised learning, we developed a novel self-generated learning model of automatic model selection with a hybrid learning strategy [20], [21]. The basic idea is as follows: the input space is partitioned hierarchically into several appropriate portions of overlapping in a discriminant learning way so that the problem complexity on a specific portion roughly matches the problem-solving capability of an available learning model, then the learning model works on its own portion of the input space [20], [21]. For constructive learning, we developed the growing and the credit-assignment algorithms for automatic model selection. As a result, the resultant tree-structured learning model would automatically assign an unknown dynamic pattern to a handful of adjacent local learning models responsible for the overlapping portions containing this pattern. It turns out that the soft partitioning of the input space enables our model to manage both intra- and inter-class variability very well, as demonstrated by a range of dynamic pattern classification tasks [21]. In [22], one can find how our learning model works and yields favorite results for a real world dynamic pattern classification problem in detail.

For dealing with the intra-class variability of dynamic patterns, many researchers adopt a model-based methodology where dynamic patterns in the same class are characterized by a set of non-parametric prototypes or a probabilistic parametric model. In terms of dynamic pattern recognition, however, such modeling has a weak discriminating capability since no

inter-class variability can be explicitly taken into account for classification. In order to overcome this weakness, we initially suggested a robust decision rule [23] that results in an improvement in discrimination. However, the poor discriminating weakness is still not overcome entirely. Our further endeavor was developing a hybrid learning model that combines a learning model of the strong discriminating capability with those models of characterizing dynamic patterns of the same class respectively [24]. As illustrated in Fig. 2, a generative model, e.g., a Gaussian mixture model [24], is used to characterize the intra-class variability of dynamic patterns belonging to the same class, whilst a supervised learning model, e.g., a neural network [24], is employed to capture the inter-class variability. Such a hybrid learning model can also be viewed as an innovative use of the divide-and-conquer principle; generative models deal with sub-problems of characterizing dynamic patterns of the same class while a learning model globally looks at all solutions of sub-problems by making use of the inter-class information for discrimination. An application in speaker identification manifests that our hybrid learning model effectively deals with intra- and inter-class variability and therefore generates the significantly better performance [24].

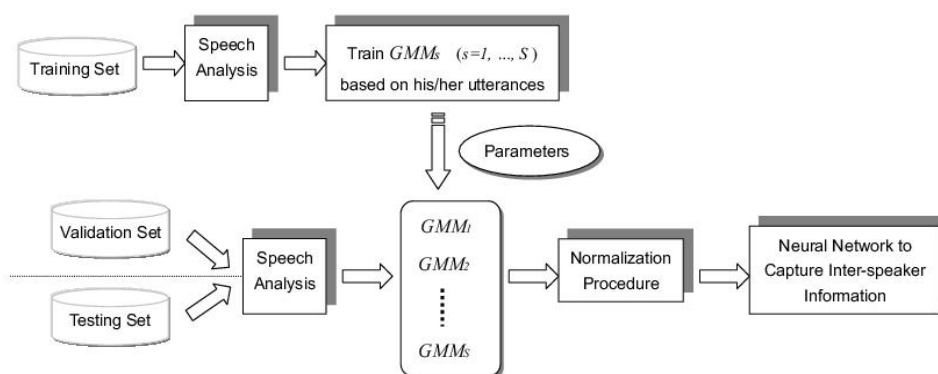


Fig. 2. A hybrid pattern recognition scheme to capture both intra- and inter-class variability as exemplified to be applied in speaker recognition [24].

In summary, our research suggests that the divide-and-conquer principle offers us a powerful tool for dynamic pattern recognition and the innovative use of this principle has resulted in the favorite performance.

2.3. On Exploitation of Intrinsic Contextual Information

For dynamic patterns, our research reveals that contrary to appearances, there is very useful intrinsic contextual information that encodes their essential structures, mutual constraints and regularities of variations of dynamic patterns. In our research, we have endeavored to explore and exploit intrinsic contextual information underlying different kinds of dynamic patterns in context of machine perception.

Most of biometric tokens fall into the category of dynamic patterns. For biometric authentication, decision-making is one of the most important issues. For decision-making on dynamic patterns, a decision rule or a threshold estimated from only training data often causes the poor generalization in testing or production data unobserved during training due to their mismatch. For good generalization, production data should be taken into account as well for the decision rule generation. Although production data are not available during training, our research unveils that there are often intrinsic contextual constraints because both training and production data belonging to the same class are generated with the same regularity. Using the uncovered regularity, we proposed a method to rectify a decision rule generated from only training data for better generalization [25]. As exemplified in speaker verification, we uncover that scores produced by a speaker model can be divided into normal

and abnormal; the former conveys useful information while the latter mainly caused by noise and other factors misleads the decision rule generation. Using this intrinsic contextual information, we developed an algorithm for pruning abnormal scores, and this algorithm readily improves the generalization [25]. Here we emphasize that unlike the cross-validation techniques, the improved generalization benefits from the exploration and the exploitation of intrinsic contextual information without use of additional production data.

In computer vision, dynamic patterns are ubiquitous, ranging from low-level image signals to high-level organizational/perceptual prototypes. As another major research topic, we have also researched into the exploitation of intrinsic contextual information in early vision [26], [30]-[32] and 3-D object recognition [33].

Early vision plays a prominent role in visual information processing, which critically determines the ultimate interpretation of an image. Due to high complexity of the real world and noise corruption from various sources during imaging, there are enormous amount of ambiguities for an early vision task. Fortunately, there is intrinsic contextual information underlying an image that can resolve ambiguities. In our recent work [26], we uncovered that unlike local discontinuities that can confuse noise with not-trivial features, there are relatively global yet robust contextual discontinuities that would robustly specify intrinsic structures of an image. As a result, we derived a contextual discontinuity measure from the scale-based affinity theory [27]. The measure detects the intrinsic contextual information of much fewer ambiguities so as to form a "road map" of silent features underlying an image [26].

Smoothing is viewed as a general tool in the low-level of machine vision for noise removal and irrelevant feature elimination demanded by subsequent processing. By combining the local and contextual discontinuities in a synergetic way, we proposed a novel adaptive smoothing [26]. In general, there are two weaknesses for an adaptive smoothing algorithm; one is that its performance is extremely sensitive to the termination time, and the other is that it often fails to eliminate the impulse noise. The use of contextual discontinuities readily overcomes the two weaknesses and yields the favorite outcome; our adaptive smoothing scheme is, to a great extent, immune to termination times so that non-trivial features can be preserved in a widely range of iterations, and our algorithm can remove the impulse noise mixed with other types of noise (for details and a formal analysis, see [26]).

As a major aspect of visual perception, image segmentation is a process that partitions an image into a set of coherent regions. Recently, a biologically plausible computational model [28], *locally excitatory and globally inhibitory oscillator networks* (LEGION), was proposed for perceptual modeling and has been successfully applied to image segmentation [29]. Without the use of any intrinsic contextual information, the original LEGION algorithm [29] fails to work in the presence of substantial noise [30], [31]. In our research, we extended the LEGION by exploiting intrinsic neighboring constraints [30] and introducing a contextual discontinuity map [31], which results in two improved variants of the LEGION. In [30], we introduce a set of neighborhoods to generate dynamic coupling structures associated with a specific oscillator. As a result, we used an ensemble of oscillators dynamically coupled in a local region instead of pair-wise coupling and further developed two novel grouping rules for perceptual organization [30]. Owing to the use of intrinsic neighboring structural constraints, the resultant dynamic coupling without iterative operations makes our segmentation network robust to noise in an image, as demonstrated with a variety of images [30]. On the other hand, the use of a contextual discontinuity map introduces an additional weight adaptation mechanism to the LEGION [31]. Such a weight adaptation mechanism tends to remove noise and preserve significant discontinuities in an image prior to segmentation and its usefulness has been demonstrated by a range of synthetic and real image segmentation tasks [31]. Furthermore, we have also applied this LEGION variant of weight adaptation to a challenging real world problem [32], namely extraction of hydrographic objects from satellite images, which yields the favorite extraction results. As illustrated in Fig. 3, the river in Fig.

3(a) is extracted and shown in Fig. 3(b) by marking it as white and superimposing it on the original image. It is obvious that all boundaries are located accurately even for those of small islands covered by forests and the very narrow bridge. Fig. 3(c) shows the corresponding part of the topographic map. By comparing Fig. 3(a) and 3(c), one can see that hydrographic objects have changed from the map. Fortunately, our system precisely detects this change and therefore is suitable for map revision. Furthermore, Fig. 3(e) shows the extraction result from Fig. 3(d), a satellite image of the Washington East, D.C. and Maryland area, U.S.A., where all major hydrographic objects are extracted with high accuracy (for details, see [32]).

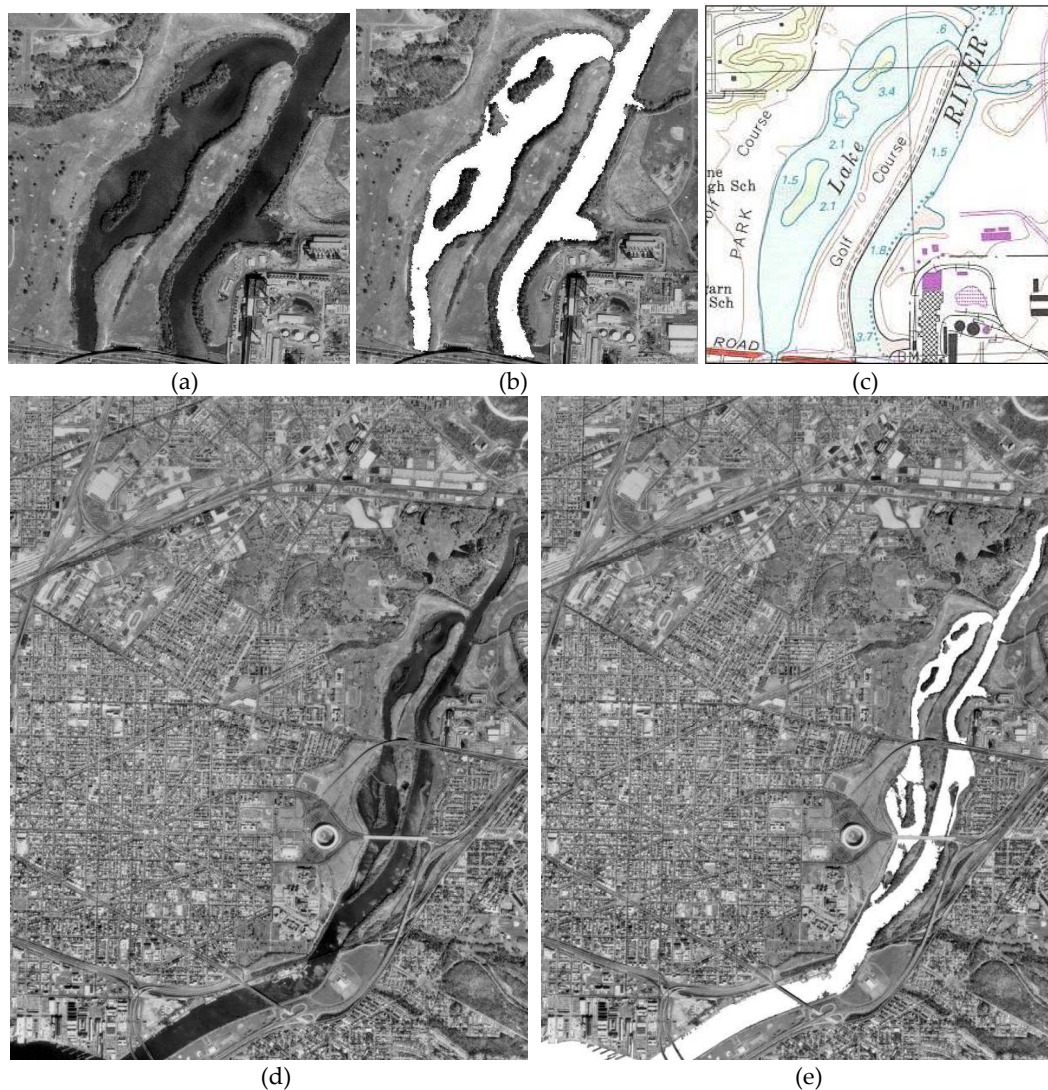


Fig. 3. Extraction of hydrographic objects by using intrinsic contextual information [26], [32]. (a) A satellite image patch of 640×640 pixels. (b) The extraction result from the patch in (a). (c) A topographic map corresponding to the patch in (a). (d) A larger satellite image of 7676×6204 pixels containing the patch in (a). (e) The extraction result from the image in (d).

3-D object recognition from a 2-D image is an ill-posed problem; multiple objects can be coincidentally projected into the same image. Thus, corrupted by noise during imaging, numberless aspects of a 3-D object can be viewed as dynamic patterns. Psychological studies in human vision system uncover that a multi-stage process takes place in 3-D object recognition via a range of perceptual organization operations for recovery from non-trivial image features to aspects. In our earlier research, we proposed a distributed representation of aspects [33] that models intrinsic or non-accidental contextual constraints from image features to aspects in a hierarchical way for perceptual organization. This distributed aspect

representation makes a parallel voting scheme feasible so that non-trivial aspects can be discovered and further recovered quickly by means of intrinsic contextual constraints [33]. This preliminary work indicates that the exploitation of intrinsic contextual information can effectively tackle a difficult problem of computer vision even in a complex environment.

In summary, our research suggests that the exploitation of intrinsic contextual information underlying dynamic patterns effectively resolves ambiguities and thus yields the robust performance for dynamic pattern analysis, discovery and recognition.

3. Ongoing Research Topics

As described in this article, dynamic patterns lead to a large number of challenging problems in different fields of artificial intelligence. As one of the long-term research interests, the author and his collaborators would continue to work for dynamic pattern analysis, discovery and recognition, including theories and their applications. At the University of Manchester, the author and his students are now researching into the following topics relevant to dynamic pattern analysis, discovery and recognition.

- **Exploration/Exploitation of representations:** On the basis of our previous research, we mainly focus on two problems on the representation of dynamic patterns. One is the exploration of a proper representation of dynamic patterns by learning to filter the irrelevant information out for a task-driven task. The other is the exploitation of different representations for dynamic patterns that cannot be well presented by a single representation. We are improving our proposed soft-competition scheme on different representations [4], [10] and shall further study an alternative representation fusion scheme of automatic feature selection. In terms of applications, we are studying the two problems on various dynamic biometric tokens for personal authentication and dynamic patterns in video for automatic content analysis.
- **Task-driven similarity measures:** Since dynamic patterns often convey mixing information, a generic similarity criterion may fail to work in accurately measuring the similarity between dynamic patterns for a specific task. In our ongoing research, we are developing an approach for finding a task-driven similarity measure by learning from selected data with a hybrid learning strategy. The proposed approach will be applied to speaker clustering, meaningful audio stream segmentation from video and affective computing, e.g., facial expression analysis and recognition.
- **Hybrid learning paradigm:** As teacher's information is available for a task, one often takes it for granted that the supervised learning paradigm should be simply used. For a complex dynamic pattern recognition task, however, our research unveils that the use of only a supervised learning model is often very difficult to achieve a solution as expected even though teacher's information is available. In our previous research [20], [21], we have successfully combined unsupervised and learning paradigms to develop a self-generated hybrid learning model by the innovative use of the divide-and-conquer principle for supervised learning. However, partitioning the input space was based on a heuristic algorithm [20], which does not guarantee an optimal solution in terms of generalization though the hybrid learning model yields the satisfactory performance for a range of dynamic pattern recognition problems. In our ongoing research, we are improving our hybrid learning model by introducing an optimization criterion to automatically partition the input space and a "compatibility" measure to examine whether a recognition sub-task specified on a specific portion of the input space can be handled by a given learner of the limited learning capability. In addition, we are going to research into other hybrid learning models (e.g., see [34] for an alternative hybrid learning model used in news reading personalization).

We anticipate that our continuous endeavors will lead to effective and easy-to-use techniques for dynamic pattern analysis, discovery and recognition, and their applications would solve challenging real world problems related to dynamic patterns.

Acknowledgment: Parts of the research reported in this article were supported by the National Science Foundation in China, the U.S. National Science Foundation, the U.S. Office of Naval Research, Microsoft Research Asia and Japan Society of the Promotion of Science. The author would also like to acknowledge contributions of all the collaborators who got involved in research projects reported in this article.

References

- [1] A. K. Jain, R. P. W. Duin and J. Mao, "Statistical pattern recognition: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, 2000.
- [2] X. D. Huang, A. Acero and H. W. Hon, *Spoken Language Processing*, New York: Wiley, 2000.
- [3] B. Xiang, K. Chen and H. Chi, "Explore perceptually-processing based features for speaker recognition," *Journal of Electronics*, vol. 6, no. 2, pp. 56-63, 1996.
- [4] K. Chen, "On the use of different speech representations for speaker modeling," *IEEE Transactions on Systems, Man, and Cybernetics*, vol.35, no. 3, pp. 301-314, 2005.
- [5] K. Chen, D. H. Xie and H. S. Chi, "Speaker identification based on hierarchical mixture of experts," In *Proceedings of World Congress on Neural Networks*, pp. 493-496, 1995.
- [6] K. Chen, D. H. Xie and H. S. Chi, "Speaker identification based on the time-delay hierarchical mixture of experts," In *Proceedings of IEEE International Conference on Neural Networks*, pp. 2062-2066, 1995.
- [7] K. Chen, D. H. Xie and H. S. Chi, "Speaker identification using the time-delay HMEs," *International Journal of Neural Systems*, vol. 7, no. 1, pp. 29-43, 1996.
- [8] K. Chen, L. Wang and H. S. Chi, "Methods of combining multiple classifiers with different feature sets and their applications to text-independent speaker identification," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 11, no. 3, pp. 417-445, 1997.
- [9] K. Chen and H. S. Chi, "A method of combining multiple probabilistic classifiers through soft competition on different feature sets," *Neurocomputing*, vol. 20, no. 1-3, pp. 227-252, 1998.
- [10] K. Chen, "A connectionist method for pattern classification with diverse feature sets," *Pattern Recognition Letters*, vol. 19, no. 7, pp. 545-558, 1998.
- [11] K. Chen and H. S. Chi, "A modular neural network architecture for pattern classification with different feature sets," *International Journal of Neural Systems*, vol. 9, no. 6, pp. 563-581, 1999.
- [12] R. I. Jacobs, M. I. Jordan, S. Nowlan and G. Hinton, "Adaptive mixture of local experts," *Neural Computation*, vol. 3, no. 1, pp. 79-87, 1991.
- [13] M. I. Jordan and R. I. Jacobs, "Hierarchical mixture of experts and the EM algorithm," *Neural Computation*, vol. 6, no. 2, pp. 181-214, 1994.
- [14] Y. Bengio and P. Frasconi, "Input-output HMMs for sequence processing," *IEEE Transactions on Neural Networks*, vol. 7, no. 5, pp. 1231-1249, 1996.
- [15] K. Chen, D. H. Xie and H. S. Chi, "Text-dependent speaker identification based on input/output HMMs," *Neural Processing Letters*, vol. 3, no. 2, pp. 81-89, 1996.
- [16] K. Chen, D. H. Xie and H. S. Chi, "A modified HME architecture for text-dependent speaker identification," *IEEE Transactions on Neural Networks*, vol. 7, no. 5, pp. 1309-1313, 1996.
- [17] K. Chen, L. Xu and H. S. Chi, "Improved learning algorithms for mixture of experts in multiclass classification," *Neural Networks*, vol. 12, no. 9, pp. 1229-1252, 1999.
- [18] R. E. Schapire, "The boosting approach to machine learning: An overview," *Proceedings of MSRI Workshop on Nonlinear Estimation and Classification*, pp. 1-23, 2002.
- [19] K. Chen, "Boosting input/output hidden Markov models for sequence classification," In *Lecture Note in Computer Science 3611: Proceedings of International Conference on Natural Computation*, Springer-Verlag, pp. 656-665, 2005.
- [20] K. Chen, X. Yu and H. S. Chi, "Combining linear discriminant functions with neural networks for supervised learning," *Neural Computing & Applications*, vol. 6, no. 1, pp. 19-41, 1997.

- [21] K. Chen, L. P. Yang, X. Yu and H. S. Chi, "A self-generated modular neural network architecture for supervised learning," *Neurocomputing*, vol. 16, no. 1, pp. 33-48, 1997.
- [22] K. Chen, X. Yu and H. S. Chi, "Text-dependent speaker identification based on the modular tree," *Journal of Electronics*, vol. 5, no. 1, pp. 63-69, 1996.
- [23] K. Chen, T. Y. Wu and H. J. Zhang, "On the use of nearest feature lines for speaker identification," *Pattern Recognition Letters*, vol. 23, no. 14, pp. 1735-1746, 2002.
- [24] L. Wang, K. Chen and H. S. Chi, "Capture inter-speaker information with a neural network for speaker identification," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 436-445, 2002.
- [25] K. Chen, "Towards better making a decision in speaker verification," *Pattern Recognition*, vol. 36, no. 2, pp. 329-346, 2003.
- [26] K. Chen, "Adaptive smoothing via contextual and local discontinuities," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1552-1567, 2005.
- [27] P. K. Saha, J. K. Udupa and D. Odhner, "Scale-based fuzzy connected image segmentation: Theory, algorithms, and validation," *Computer Vision and Image Understanding*, vol. 77, no. 2, pp. 145-174, 2000.
- [28] D. L. Wang and D. Terman, "Locally excitatory globally inhibitory oscillator networks," *IEEE Transactions on Neural Networks*, vol. 6, no. 2, pp. 283-286, 1995.
- [29] D. L. Wang and D. Terman, "Image segmentation based on oscillatory correlation," *Neural Computation*, vol. 9, no. 3, pp. 805-836, 1997.
- [30] K. Chen and D. L. Wang, "A dynamically coupled neural oscillator network for image segmentation," *Neural Networks*, vol. 15, no. 3, pp. 423-439, 2002.
- [31] K. Chen, D. L. Wang and X. W. Liu, "Weight adaptation and oscillatory correlation for image segmentation," *IEEE Transactions on Neural Networks*, vol. 11, no. 5, pp. 1106-1123, 2000.
- [32] X. W. Liu, K. Chen and D. L. Wang, "Extraction of hydrographic regions from remote sensing images using an oscillator network with weight adaptation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 207-211, 2001.
- [33] K. Chen and M. Ishikawa, "A parallel voting scheme for aspect recovery," *Journal of Computer Science and Technology*, vol. 10, no. 5, pp. 385-402, 1995.
- [34] K. Chen and S. Yeung, "Personalized news reading via hybrid learning," In *Lecture Note in Computer Science 3177: Proceedings of International Conference on Intelligent Data Engineering and Automatic Learning*, Springer-Verlag, pp. 607-612, 2004.



Ke Chen (M'97-SM'00) received the BS and MS degrees from Nanjing University in 1984 and 1987, respectively, and the PhD degree from Harbin Institute of Technology in 1990, all in computer science. He has been with The University of Manchester since 2003, where he is a senior lecturer in computer science. He was with The University of Birmingham, Peking University, The Ohio State University, Kyushu Institute of Technology and Tsinghua University. He was a visiting professor at Microsoft Research Asia and Hong Kong Polytechnic University. He serves as an associate editor of the *IEEE Transactions on Neural Networks* and the *International Journal of Image and Graphics*, a program co-chair of

ICNC'05, and has been a member of technical program committee of numerous international conferences including IJCNN'03 and CogSci'06. At present, he chairs the audio & speech processing task force of the Neural Network Technical Committee and the multimedia & biometrics task force of Emergent Technology Technical Committee of IEEE Computational Intelligence Society, and is also the vice chair of IEEE Computational Intelligence Society UKRI Chapter. He is a recipient of several academic awards including the NSFC Distinguished Principal Young Investigator Award and the JSPS Research Award. His current research interests include pattern recognition, machine learning and machine perception.