ICFHR 2010 Tutorial

Markov Models for Handwriting Recognition (Duration: 3 Hrs.)

Speakers
Gernot A. Fink
TU Dortmund, Department of Computer Science,
44221 Dortmund, Germany

Email: Gernot.Fink@udo.edu

Gernot A. Fink received the diploma in computer science from the University of Erlangen-N¨urnberg, Erlangen, Germany, in 1991 and the Ph.D. degree (Dr.-Ing.) also in computer science from 3 Bielefeld University, Germany, in 1995. In 2002, he received the venia legendi (Habilitation) in Applied Computer Science from Bielefeld University. From 1991 to 2005 he was with the Applied Computer Science Group at the Faculty of Technology of Bielefeld University. Since 2005 he is professor for Pattern Recognition in Embedded Systems within the Department of Computer Science at TU Dortmund, Dortmund, Germany, where he also heads the Intelligent Systems group at the Robotics Research Institute (IRF).

His research interests lie in the development and application of pattern recognition methods in the fields of man machine interaction, multimodal machine perception including speech and image processing, statistical pattern recognition, and handwriting recognition.

Gernot Fink has published extensively on the use of Markov-Model based techniques for pattern recognition problems including a textbook [1] on the subject. He has worked successfully on problems of handwriting recognition [2, 3, 4, 8, 13] with a focus on camera-based recognition systems [7, 6, 11, 12, 14]. Gernot Fink has a long record of teaching in the field of pattern recognition (lectures on speech recognition and image processing) and of more general topics in computer science (computer architecture and software engineering). He has given invited talks on many different topics from his research interests on various occasions. Additionally, together with Thomas Pl¨otz he has given a successful tutorial on Markov Models for Handwriting Recognition at ICDAR 2007.

Topics to be Covered

1. Motivation

Starting from the classical architecture of character recognition systems that apply the segmentation and classification paradigm the advantages of a statistical framework allowing for implicit segmentation will be motivated. The application areas and prototypical system architectures for the relevant cases of online and offline recognition of handwriting will be presented. For offline recognition the principle idea behind the representation of image data as sequences will be described.

2. Theoretical Concepts

The second part of the tutorial starts by introducing the statistical framework of pattern recognition as applied to sequence data. Then the theoretical concepts behind Hidden Markov Models (HMMs) and so-called language models realized as Markov chain models will be presented (cf. [1]). The following topics will be covered:

– Hidden Markov Models
– Definition
– Modeling of Output Probability Density Functions (with emphasis on continuous outputs, mixture densities)

– Usage Concepts (how a generative statistical model can be applied for pattern recognition)
– Scoring (determining how well a model describes sample data, i.e., computing the generation probability [Forward-Algorithm])
– Decoding (using structured models for recognition and implicit segmentation of data [Viterbi-Algorithm])
– Parameter Estimation (basic concepts of estimating model parameters from sample data [Forward-Backward-Algorithm] and training algorithms [Baum-Welch-Algorithm])
– *n*-Gram Language Models
– Definition
– Usage Concepts (what to describe with a Markov chain model?)
– Scoring (determining how well a language model can predict unseen data introducing the notion of perplexity)
– Parameter Estimation (methods for robust estimation of model parameters in the presence of unseen events: Discounting, Interpolation and Backing-Off)

3. Practical Aspects

Markov models cannot be used successfully in practice given their theoretical foundations only. Therefore, in the third part of the tutorial, the most important concepts, methods, and algorithms relevant in practical applications of the Markov model paradigm will be described.

– Computations with Probabilities (logarithmic representation, flooring)
– Configuration of HMMs (model topologies, model construction from basic units, compound models)
– Robust Parameter Estimation (parameter tying, especially mixture tying; parameter initialization [vector quantization and segmental *k*-means])
– Efficient Model Evaluation (beam search; forward-backward-pruning; tree-based representation of large lexica)
– Integrated Search (effective combinations of HMMs and *n*-Gram models via specialized integrated search algorithms)

4. Putting It All Together

In order to give the audience a good impression of how the concepts described are combined in real systems we will present in detail the architecture and processing steps of our own system for the recognition of unconstrained offline handwritten text (cf. [11, 14]) and relate those solutions to the ones found in similar systems reported in the literature (cf. [10]). In addition to the actual configuration of the statistical models used, the methods applied for pre-processing and normalization, for feature extraction, and for integrated decoding will be presented.

References

[1] G. A. Fink. Markov Models for Pattern Recognition, From Theory to Applications. Springer, Heidelberg, 2008.

[2] G. A. Fink and T. Pl¨otz. On appearance-based feature extraction methods for writer-independent handwritten text recognition. In Proc. Int. Conf. on Document Analysis and Recognition, volume 2, pages 1070–1074, Seoul, Korea, 2005. IEEE.

[3] G. A. Fink and T. Pl¨otz. Unsupervised estimation of writing style models for improved unconstrained off-line handwriting recognition. In Proc. Int. Workshop on Frontiers in Handwriting Recognition, pages 429–434, La Baule, France, 2006.

[4] G. A. Fink and T. Pl¨otz. On the use of context-dependent modelling units for HMM-based offline handwriting recognition. In Proc. Int. Conf. on Document Analysis and Recognition, pages 729–733, Curitiba, Brazil, 2007.

[5] G. A. Fink and T. Pl¨otz. Developing pattern recognition systems based on Markov models: The ESMERALDA framework. Pattern Recognition and Image Analysis, 18(2):207–215, June 2008.

[6] G. A. Fink and M. Wienecke. Experiments in video-based whiteboard reading. In First Int. Workshop on Camera-Based Document Analysis and Recognition, pages 95–100, Seoul, Korea, 2005.

[7] G. A. Fink, M. Wienecke, and G. Sagerer. Video-based on-line handwriting recognition. In Proc. Int. Conf. on Document Analysis and Recognition, pages 226–230, Seattle, 2001. IEEE.
[8] N. Y. Hammerla, T. Pl¨otz, S. Vajda, and G. A. Fink. Towards feature learning for HMM-based offline handwriting recognition. In International Workshop on Frontiers of Arabic Handwriting Recognition, Istanbul, Turkey, 2010. IAPR. (accepted).
[9] U.-V. Marti and H. Bunke. The IAM-database: An English sentence database for offline handwriting recognition. Int. Journal on Document Analysis and Recognition, 5(1):39–46, 2002.
[10] T. Pl¨otz and G. A. Fink. Markov Models for Offline Handwriting Recognition: A Survey. Int. Journal on Document Analysis and Recognition, 12(4):269–298, 2009.
[11] T. Pl¨otz, C. Thurau, and G. A. Fink. Camera-based whiteboard reading: New approaches to a challenging task. In Proc. 11th Int. Conference on Frontiers in Handwriting Recognition, pages 385–390, Montreal, Canada, 2008.