

# MARF : Modular Audio Recognition Framework

Serguei A. Mokhov

Département d'Informatique et Génie Logiciel  
Faculté de Génie et Informatique  
Université Concordia, Montréal, Québec, Canada  
[mokhov@cse.concordia.ca](mailto:mokhov@cse.concordia.ca)

ACFAS 2010

## Introduction

Résumé

Le pipeline classique

Applications de MARF

## Quelques exemples résultats

Résultats de SpeakerIdentApp du MARF, qui ?

Résultats de SpeakerIdentApp du MARF, sexe ?

Résultats de SpeakerIdentApp du MARF, accent ?

## Conclusion

Les travaux futurs

Remerciements

Références

# Introduction I

- ▶ Le Modular Audio Recognition Framework (MARF) [13, 1, 5, 4, 12] conçu en 2002, est une plateforme de recherche open-source et une collection de composants avec des algorithmes pour le traitement de la voix, le son, la parole, et l'écriture et de langues naturelles (TALN) statistique [7].

## Introduction II

- ▶ MARF a été créée en Java et organisée sous forme de modules extensibles qui facilitent l'ajout de nouveaux algorithmes.
- ▶ MARF peut être utilisée comme une bibliothèque dans une application ou comme une base de support à l'apprentissage et en extension.

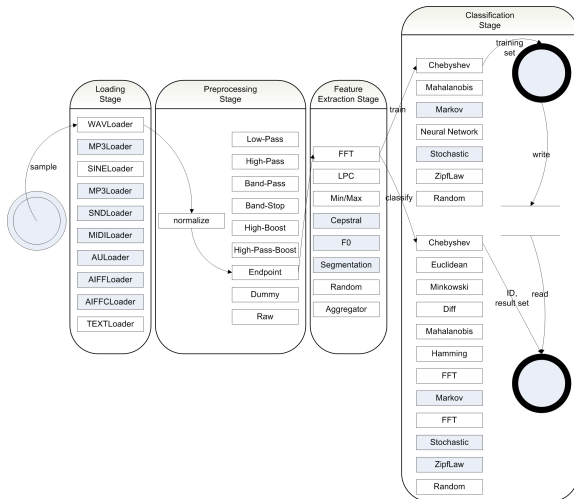
## Introduction III

- ▶ MARF a aussi été publié dans les plusieurs articles de conférence avec les details scientifiques dedant.
- ▶ De la documentation détaillée et la référence d'API en format javadoc sont disponibles étant donné que le projet tente d'être bien-documenté.
- ▶ MARF et ses applications sont déployés sous une licence BSD.

# Etapes

1. Chargement de données (fichiers)
2. Préprocesseurs
3. Extraire de caractéristiques
4. Classifieurs (entraînement et classification)

# Le pipeline classique



# Applications de MARF

Les applications suivantes sont fournies pour montrer les utilisations possibles du framework. Les exemples d'applications incluses démontrent les aspects généraux de MARF appliqués dans plusieurs domaines au-delà de l'audio et de TALN, tel que :

- ▶ le génie logiciel [3],
- ▶ les calculs distribué [2, 10, 11] et autonome [16, 9, 15],
- ▶ le traitement multimédia,
- ▶ l'analyse forensique de fichiers [8],
- ▶ la programmation intensionnelle [6],
- ▶ la reconnaissance des auteurs, des langues, des sexes, de l'âge [4, 5],
- ▶ ainsi que la reconnaissance des formes en générale.



# Résultats de SpeakerIdentApp du MARF, qui ?

## ► Personnes parlantes (Qui ?)

Configuration	Bon	Mal	Taux de precision%
-silence -bandstop -aggr -cos	29	3	90.62
-silence -bandstop -fft -cos	29	3	90.62
-bandstop -fft -cos	28	4	87.50
-silence -noise -bandstop -fft -cos	28	4	87.50
-silence -low -aggr -cos	28	4	87.50
-silence -noise -norm -aggr -cos	28	4	87.50
-silence -low -fft -cos	28	4	87.50
-silence -noise -norm -fft -cos	28	4	87.50
-silence -noise -low -aggr -cos	28	4	87.50
-silence -noise -low -fft -cos	28	4	87.50
-bandstop -aggr -cos	28	4	87.50
-norm -fft -cos	28	4	87.50
-silence -raw -aggr -cos	28	4	87.50
-silence -noise -raw -aggr -cos	28	4	87.50
-norm -aggr -cos	28	4	87.50
-silence -noise -bandstop -aggr -cos	28	4	87.50
-silence -norm -fft -cos	27	5	84.38
-silence -norm -aggr -cos	27	5	84.38
-low -fft -cos	27	5	84.38
-noise -bandstop -aggr -cos	27	5	84.38

# Résultats de SpeakerIdentApp du MARF, sexe ?

## ► Personnes parlantes (Sexe ?)

Configuration	Bon	Mal	Taux de precision	2ieme bon	2ieme mal	2ieme taux%
-noise -high -aggr -mink	26	6	81.25	32	0	100
-silence -noise -band -aggr -cheb	26	6	81.25	32	0	100
-silence -noise -band -lpc -cos	26	6	81.25	31	1	96.88
-silence -noise -band -fft -cheb	26	6	81.25	32	0	100
-noise -bandstop -fft -diff	26	6	81.25	32	0	100
-noise -bandstop -fft -cheb	26	6	81.25	32	0	100
-silence -band -lpc -cos	25	7	78.12	31	1	96.88
-silence -noise -bandstop -fft -diff	25	7	78.12	32	0	100
-noise -endp -lpc -eucl	25	7	78.12	31	1	96.88
-silence -noise -band -aggr -eucl	25	7	78.12	32	0	100
-silence -noise -endp -lpc -cheb	25	7	78.12	32	0	100
-noise -endp -lpc -diff	25	7	78.12	32	0	100
-silence -noise -band -fft -eucl	25	7	78.12	32	0	100
-silence -noise -band -aggr -diff	25	7	78.12	32	0	100
-silence -noise -bandstop -fft -cheb	25	7	78.12	32	0	100
-silence -noise -band -fft -diff	25	7	78.12	32	0	100
-noise -bandstop -aggr -cheb	25	7	78.12	32	0	100
-noise -band -aggr -cheb	24	8	75	32	0	100
-noise -high -fft -eucl	24	8	75	31	1	96.88
-noise -high -lpc -cos	24	8	75	30	2	93.75

# Résultats de SpeakerIdentApp du MARF, accent ?

## ► Personnes parlantes (Accent ?)

Configuration	Bon	Mal	Taux de precision	2ieme bon	2ieme mal	2ieme taux%
-silence -endp -lpc -cheb	24	8	75	26	6	81.25
-bandstop -fft -cos	23	9	71.88	27	5	84.38
-low -aggr -cos	23	9	71.88	26	6	81.25
-noise -norm -aggr -cos	23	9	71.88	26	6	81.25
-noise -low -aggr -cos	23	9	71.88	26	6	81.25
-noise -bandstop -aggr -cos	22	10	68.75	27	5	84.38
-noise -low -fft -cos	22	10	68.75	26	6	81.25
-noise -bandstop -fft -cos	22	10	68.75	27	5	84.38
-norm -aggr -cos	22	10	68.75	26	6	81.25
-endp -lpc -cheb	21	11	65.62	24	8	75
-silence -noise -low -aggr -cos	21	11	65.62	25	7	78.12
-low -fft -cos	21	11	65.62	27	5	84.38
-noise -norm -fft -cos	21	11	65.62	27	5	84.38
-silence -bandstop -aggr -cos	20	12	62.5	25	7	78.12
-silence -low -aggr -cos	20	12	62.5	25	7	78.12
-silence -noise -norm -aggr -cos	20	12	62.5	25	7	78.12
-silence -bandstop -fft -cos	20	12	62.5	25	7	78.12
-silence -low -fft -cos	20	12	62.5	25	7	78.12
-silence -noise -norm -fft -cos	20	12	62.5	25	7	78.12
-silence -noise -low -fft -cos	20	12	62.5	25	7	78.12

# Conclusion

- ▶ Les résultats de note approche sont encourageants et s'appliquent au plusieurs domaines.
- ▶ A la connaissance de l'auteur, il n'y a pas d'un autre framework comme MARF pour la recherche comparative des algorithmes et leurs combinaisons pour la reconnaissance de formes et TALN.
- ▶ Le project le plus proche open-source est CMU Sphinx [14] qui vous offre la reconnaissance du parole, mais il est trop compliqué pour les experimentations que MARF peut faire. De plus, Sphinx n'offre pas de modules TALN.
- ▶ Donc, la simplicité et généralité de MARF sont les meilleures avantages. Un désavantage de MARF en particulier, c'est aussi l'approche framework nécessité de changement de framework et tous les modules qui on en depend si on ajout un nouvelle fonction ou tache de reconnaissance.

# Les travaux futurs

- ▶ Conférences TALN 2010 et DEFT 2010 du 19 au 23 juillet 2010 ici a Montréal, à la Polytechnique.
  - ▶ L'atelier DEFT 2010  
<http://www.groupe.polymtl.ca/taln2010/deft.php>  
(Atelier d'évaluation en fouille de textes sur l'identification de la période et du lieu de publication d'articles de presse francophone.)
  - ▶ Démonstration du MARF prévue dans le cadre de la conférence jointe TALN 2010  
[http://www.groupe.polymtl.ca/taln2010/appel\\_taln\\_demos.php](http://www.groupe.polymtl.ca/taln2010/appel_taln_demos.php)
- ▶ Amélioration du framework.
- ▶ Publication de nouveaux résultats aux articles journaux.

# Remerciements

Nous tenons à remercier chaleureusement tous ceux et celles qui ont contribué de près ou de loin à la réalisation de ce travail, ils se reconnaîtront.

- ▶ A vous tous pour avoir assister à cette presentation.
- ▶ Co-créateurs origineaux du MARF : Stephen Sinclair, Ian Clément, Dimitrios Nicolacopoulos, et Contributeurs du MARF R&D Group.
- ▶ Miao Song, Dr. Ching Y. Suen, Dr. Sabine Bergler, Dr. Leila Kosseim, Dr. Joey Paquet, Dr. Mourad Debbabi, Dr. Lingyu Wang, Dr. Amr M. Youssef, Dr. Emil Vassev, Ai-Thanh Nguyen, Michelle Khalifé.
- ▶ CRSNG et la Faculté de Génie et Informatique, Université Concordia.

Merci :-)

# Références I

- [1] Serguei Mokhov, Ian Clement, Stephen Sinclair, and Dimitrios Nicolacopoulos. Modular Audio Recognition Framework. Department of Computer Science and Software Engineering, Concordia University, Montreal, Canada, 2002–2003. Project report, <http://marf.sf.net>, last viewed April 2012.
- [2] Serguei A. Mokhov. On design and implementation of distributed modular audio recognition framework : Requirements and specification design document. [online], August 2006. Project report, <http://arxiv.org/abs/0905.2459>, last viewed April 2012.
- [3] Serguei A. Mokhov. Introducing MARF : a modular audio recognition framework and its applications for scientific and software engineering research. In *Advances in Computer and Information Sciences and Engineering*, pages 473–478, University of Bridgeport, U.S.A., December 2007. Springer Netherlands. ISBN 978-1-4020-8740-0. doi : 10.1007/978-1-4020-8741-7. Proceedings of CISSE/SCSS'07.
- [4] Serguei A. Mokhov. Choosing best algorithm combinations for speech processing tasks in machine learning using MARF. In Sabine Bergler, editor, *Proceedings of the 21st Canadian AI'08*, LNAI 5032, pages 216–221, Berlin Heidelberg, May 2008. Springer-Verlag. doi : 10.1007/978-3-540-68825-9\\_21.



## Références II

- [5] Serguei A. Mokhov. Study of best algorithm combinations for speech processing tasks in machine learning using median vs. mean clusters in MARF. In Bipin C. Desai, editor, *Proceedings of C3S2E'08*, pages 29–43, Montreal, Quebec, Canada, May 2008. ACM. ISBN 978-1-60558-101-9. doi : 10.1145/1370256.1370262.
- [6] Serguei A. Mokhov. Towards syntax and semantics of hierarchical contexts in multimedia processing applications using MARFL. In *Proceedings of the 32nd Annual IEEE International Computer Software and Applications Conference (COMPSAC)*, pages 1288–1294, Turku, Finland, July 2008. IEEE Computer Society. doi : 10.1109/COMPSAC.2008.206.
- [7] Serguei A. Mokhov. Evolution of MARF and its NLP framework. In *Proceedings of C3S2E'10*, pages 118–122. ACM, May 2010. ISBN 978-1-60558-901-5. doi : 10.1145/1822327.1822344.
- [8] Serguei A. Mokhov and Mourad Debbabi. File type analysis using signal processing techniques and machine learning vs. `file` unix utility for forensic analysis. In Oliver Goebel, Sandra Frings, Detlef Guenther, Jens Nedon, and Dirk Schadt, editors, *Proceedings of the IT Incident Management and IT Forensics (IMF'08)*, LN140, pages 73–85. GI, September 2008. ISBN 978-3-88579-234-5.

## Références III

- [9] Serguei A. Mokhov and Emil Vassev. Autonomic specification of self-protection for Distributed MARF with ASSL. In *Proceedings of C3S2E'09*, pages 175–183, New York, NY, USA, May 2009. ACM. ISBN 978-1-60558-401-0. doi : 10.1145/1557626.1557655.
- [10] Serguei A. Mokhov, Lee Wei Huynh, and Jian Li. Managing distributed MARF with SNMP. Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada, April 2007. Project report. Hosted at <http://marf.sf.net> and <http://arxiv.org/abs/0906.0065>, last viewed February 2011.
- [11] Serguei A. Mokhov, Lee Wei Huynh, and Jian Li. Managing distributed MARF's nodes with SNMP. In *Proceedings of PDPTA'2008*, volume II, pages 948–954, Las Vegas, USA, July 2008. CSREA Press. ISBN 1-60132-082-5.
- [12] Serguei A. Mokhov, Miao Song, and Ching Y. Suen. Writer identification using inexpensive signal processing techniques. In Tarek Sobh and Khaled Elleithy, editors, *Innovations in Computing Sciences and Software Engineering ; Proceedings of CISSE'09*, pages 437–441. Springer, December 2009. doi : 10.1007/978-90-481-9112-3\\_74. ISBN : 978-90-481-9111-6, online at : <http://arxiv.org/abs/0912.5502>.

## Références IV

- [13] The MARF Research and Development Group. The Modular Audio Recognition Framework and its Applications. [online], 2002–2012. <http://marf.sf.net> and <http://arxiv.org/abs/0905.1235>, last viewed April 2012.
- [14] The Sphinx Group at Carnegie Mellon. The CMU Sphinx group open source speech recognition engines. [online], 2007–2012. <http://cmusphinx.sourceforge.net>.
- [15] Emil Vassev and Serguei A. Mokhov. Self-optimization property in autonomic specification of Distributed MARF with ASSL. In Boris Shishkov, Jose Cordeiro, and Alpesh Ranchordas, editors, *Proceedings of ICSSOFT'09*, volume 1, pages 331–335, Sofia, Bulgaria, July 2009. INSTICC Press. ISBN 978-989-674-009-2. doi : 10.5220/0002257303310335.
- [16] Emil Vassev and Serguei A. Mokhov. Towards autonomic specification of Distributed MARF with ASSL : Self-healing. In *Proceedings of SERA 2010 (selected papers)*, volume 296 of *SCI*, pages 1–15. Springer, 2010. ISBN 978-3-642-13272-8. doi : 10.1007/978-3-642-13273-5\\_1.