



HAL
open science

Reconnaissance émotionnelle par l'analyse des expressions faciales dans un tuteur intelligent affectif

Roger Nkambou, Vincent Heritier

► **To cite this version:**

Roger Nkambou, Vincent Heritier. Reconnaissance émotionnelle par l'analyse des expressions faciales dans un tuteur intelligent affectif. Technologies de l'Information et de la Connaissance dans l'Enseignement Supérieur et l'Industrie, Oct 2004, Compiègne, France. pp.149-155. edutice-00000702

HAL Id: edutice-00000702

<https://edutice.hal.science/edutice-00000702v1>

Submitted on 15 Nov 2004

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reconnaissance émotionnelle par l'analyse des expressions faciales dans un tuteur intelligent affectif

Roger NKAMBOU & Vincent HERITIER

Laboratoire GDAC – Université du Québec à Montréal
Nkambou.roger@uqam.ca ; heritier.vincent@courrier.uqam.ca

Résumé

Dans le contexte éducatif classique, l'observation et l'identification de l'état émotionnel de l'apprenant permettent à l'enseignant d'entreprendre des actions qui vont influencer la qualité de l'apprentissage et son exécution. Afin d'assurer une meilleure interaction avec l'étudiant, un système tutoriel intelligent (STI) devra tenir compte des relations entre l'émotion, la cognition et l'action en contexte d'apprentissage en s'inspirant de modèles cognitifs et comportementaux. Un tel STI, que nous appellerons STI affectif, devra d'une part reconnaître les changements affectifs (notamment émotionnels) de l'apprenant durant une session d'apprentissage, et d'autres parts exprimer des états affectifs en réponse aux événements qui surviennent durant une session d'apprentissage. Dans cet article, nous proposons une architecture de STI affectif intégrant plusieurs agents conçus pour la gestion des émotions. Nous nous focalisons plus particulièrement sur une fonction d'extraction d'émotions à partir de l'analyse des expressions faciales de l'apprenant.

Mots-clés : reconnaissance émotionnelle, apprentissage et exécution de l'image, réseau de neurones artificiels, système tutoriel intelligent, processus d'apprentissage.

Abstract:

In the classic educational context, the observation and the identification of the learner's emotional state allow the teacher to adapt the lesson, with the aim of improving the quality of the learning process. In computer environment, Intelligent Tutoring Systems (ITS) have to consider relationships between emotion, cognition and action of the learning context. Thus, inspired by cognitive and behavioral models, ITS will be able to provide pertinent tool for efficient and pleasant learning. In order to provide ITS the capacity of considering and managing the emotions of the learner, we propose an architecture which include several agents specialized in emotion recognition, based only on facial expression analysis.

Keywords: Emotion recognition, Image acquisition and processing, Artificial neural networks, Intelligent Tutoring System, Learning Process

Introduction

Dans le contexte social, les émotions ont un rôle prépondérant aussi bien en communication orale que non-verbale. Leur perception étant multimodale, l'aptitude d'un interlocuteur à les identifier à travers une

variété de comportements tels que les mouvements du visage, les gestes, le discours, constitue une base essentielle pour l'initiation de ses propres actions et réponses car le jugement et la prise de décision sont influencés par l'humeur, les sentiments mécanismes facilitant l'adaptation humaine et l'intégration sociale. Dans une activité cognitive telle que l'apprentissage, la communication est un aspect primordial et sa qualité peut influencer l'apprentissage. Ainsi, la pertinence d'un apprentissage implique autant des aspects intellectuels que socio-émotionnels ; enseigner sous-entend observer le comportement de l'élève afin de détecter des réponses affectives qui peuvent être la manifestation de sentiments d'intérêt, d'excitation, de confusion, d'anéantissement, etc. pour influencer par la suite la qualité de l'apprentissage et son exécution.

Dans les mêmes perspectives et soucis pédagogiques, les Systèmes Tuteurs Intelligents doivent tenir compte du facteur émotionnel et tenter de le gérer. Néanmoins, avec les nouvelles technologies de l'information surgissent de nouveaux problèmes méthodologiques, notamment celui de la détection et de l'évaluation des états émotionnels. Les processus mentaux de l'apprenant ne sont pas directement observables par la machine et, jusqu'à lors, les méthodes d'évaluation des changements d'état affectif les plus objectives et les moins intrusives, utilisent des capteurs externes en vue d'examiner le maintien de la posture, le regard et les expressions faciales de l'apprenant [1 ; 2].

Afin de promouvoir une communication plus dynamique et flexible entre l'apprenant et le système, nous proposons l'intégration de deux agents émotionnels adaptatifs au sein d'un STI multi-agent. Le premier permet de gérer l'expression des émotions par le tuteur suite aux actions de l'étudiant. Un tuteur émotionnel appelé (Emilie) [3] a été intégré avec succès dans un environnement de formation en ligne pour l'apprentissage des sciences. Le second a pour rôle de capturer et de gérer les émotions exprimées par l'apprenant lors d'une session d'apprentissage et ce à travers l'expression faciale de ce dernier. Cet agent (Emilie-2) fonctionne selon une boucle pédagogique incluant les actions suivantes : la capture, l'extraction et la reconnaissance des émotions à travers l'expression faciale, l'analyse et l'interprétation des émotions reconnues et de leurs effets sur l'apprentissage, le diagnostic de ces émotions et la remédiation. Nous nous intéressons dans cet article aux fonctions d'acquisition, d'extraction et de classification des émotions de Emilie-2. Nous présentons tout d'abord le cadre contextuel de la gestion des émotions en situation d'apprentissage. Nous détaillerons ensuite l'implémentation de Emilie-2, plus particulièrement son composant de reconnaissance

d'émotions. Les résultats de ces expérimentations seront ensuite présentées et discutées. Enfin, une description des travaux similaires sera effectuée, mettant en avant notre contribution.

Cadre Contextuel

Nous nous intéressons ici au traitement émotionnel en contexte d'apprentissage. Nous pensons en effet que, pour favoriser l'apprentissage, l'enseignement doit tenir compte des états émotionnels de l'apprenant. Partant de ce principe, notre recherche, située dans un contexte de Systèmes Tuteurs Intelligents, a pour objectif le développement d'un modèle de l'apprenant étendu, prenant en compte son aspect émotionnel. La figure 1 présente l'architecture d'un STI multi-agent incluant les deux agents émotionnels indiqués plus haut. Dans cette figure, le modèle de l'apprenant représente à la fois son état affectif (connaissances, compétences et historiques des performances) et son état affectif (humeurs, émotions et profil psychologique). Nous pouvons aussi remarquer que l'état cognitif est géré par un agent spécifique : le profiler.

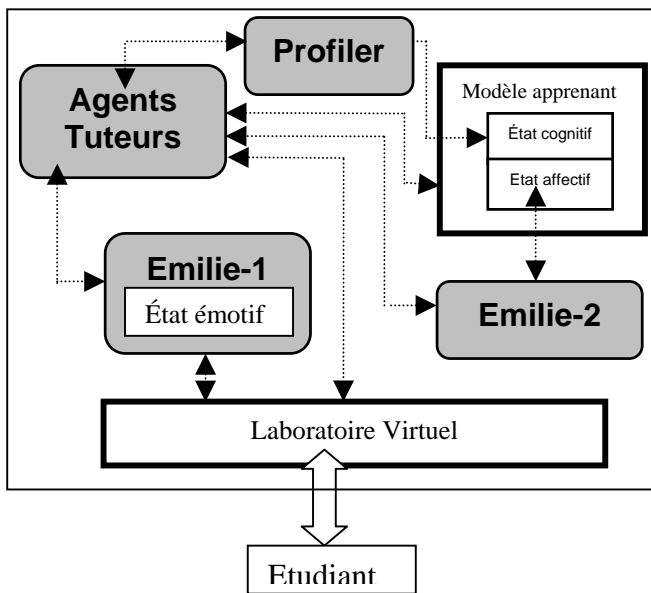


Figure 1 : Architecture générale du système

Cet agent a la responsabilité de mettre à jour les connaissances acquises, les compétences et performances de l'étudiant tout en assurant l'intégrité de son modèle cognitif. Il aide aussi au diagnostic des connaissances ou compétences mal acquises ou manquantes et permet une remédiation avec l'aide des agents tutoriels. L'état affectif contient des informations à court (resp. à moyen) et (resp. à long) terme qui correspondent aux émotions (resp. aux humeurs) et (resp. au profil psychologique) de l'apprenant. Cette partie du modèle est gérée par l'agent Emilie2. Les agents tutoriels sont des agents qui contribuent à la formation : un planificateur pour la planification de l'apprentissage, un coach pour l'aide à la réalisation de la tâche etc. Lors de l'interaction, certains agents tutoriels peuvent vouloir couvrir la dimension émotionnelle. Pour ce faire, ils doivent solliciter les services de Emilie1 qui, étant donné une activité, peut

traduire à travers un personnage (2D ou 3D) les émotions de l'agent tutoriel. Il doit au préalable être *instruit* (par le concepteur et l'agent tutoriel concerné) à la tâche concernée et aux conditions de réactions émotionnelles désirées. L'état émotionnel de Emilie1 est une mémoire à court terme qui traduit la réaction émotionnelle actuelle. L'action se passe au niveau du laboratoire virtuel qui est le seul lieu du STI accessible par l'étudiant. Ce laboratoire comprend un ensemble d'outils d'interactions (objets ou ressources) nécessaires pour l'apprentissage. Emilie-1 est décrit en détail dans [3].

Emilie-2 : Un agent de reconnaissance émotionnelle

L'Architecture de Emilie-2

Emilie2 a pour objectif, grâce à une caméra numérique placée sur l'écran de l'ordinateur, de réaliser une acquisition de l'image du visage et d'analyser les expressions faciales présentées afin d'identifier les émotions de l'étudiant en situation d'apprentissage. Son architecture comporte trois couches (figure 2) : le premier module (couche perception) extrait les expressions faciales (acquisition et tracking facial) et les catégorise (classification) ; le second (couche cognition) réalise tout d'abord une analyse de l'état émotionnel de l'apprenant. Cette analyse a pour but de permettre de traduire le sens de l'émotion dans le contexte d'apprentissage. Elle est réalisée en se basant sur plusieurs facteurs dont : l'émotion reconnue, le profil affectif actuel, l'historique des actions réalisées avant l'apparition de l'émotion, l'état cognitif, l'évolution de l'émotion, le contexte social (s'il s'agit d'un apprentissage de type social ou collaboratif).

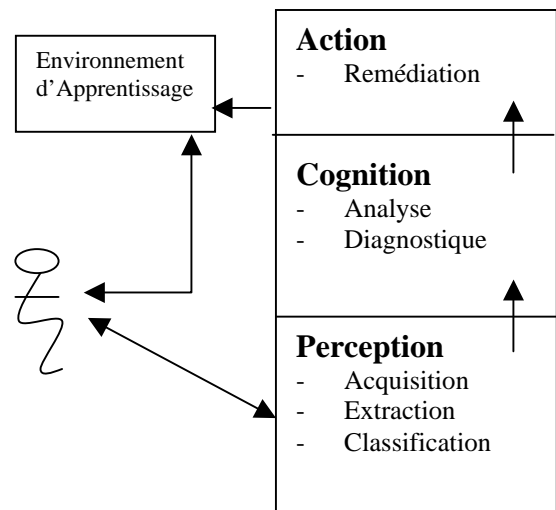


Figure 2 : Architecture de Emilie-2

Les expressions en entrée dans cette couche sont de type « joie », « peur », « dégoût », « tristesse », « colère », « surprise » et l'analyse permet de déterminer si l'étudiant éprouve de la « satisfaction », de la « confiance », de la « surprise », de la « confusion » ou encore de la « frustration ». Ces derniers sont plus

précises et permettent de prendre des actions pédagogiques appropriées à leur correction.

Un diagnostic de l'état émotionnel analysé est ensuite établi. Il déterminera ainsi les causes ayant menées à cette situation (succès/échec à un exercice, difficulté des travaux, manque de connaissances, etc.), en s'appuyant à nouveau sur l'état cognitif de l'apprenant.

La couche action consiste à définir, s'il y a lieu, un ensemble de tâches permettant de remédier à l'état émotionnel observé, afin d'amener l'étudiant dans un état plus propice à l'assimilation des connaissances.

Les couches Cognition et Action ont été implémentée avec succès dans l'agent Profiler [4 ; 5] et sont transférées dans Emilie-2 puisque nous supposons qu'un état émotionnel observé est une conséquence d'un acte d'apprentissage impliquant un ou plusieurs éléments de connaissances. Dans ces conditions, les changements émotionnels peuvent se justifier en terme cognitif. Ainsi, dans le présent article, nous nous limiterons à la présentation des détails de la couche Perception incluant les fonctions d'extraction et de classification des expressions faciales.

La couche de perception de Emilie-2

Cette couche a pour objectif la perception et la reconnaissance des émotions exprimées par l'étudiant, via une analyse de l'expression faciale. La couche de perception contient 3 étapes : l'extraction, la réduction de des données et la classification. Nous présentons dans cette section notre approche pour l'implémentation de ces fonctions dans Emilie-2.

Extraction des expressions faciales

Pour favoriser l'autonomie de l'agent, l'extraction des expressions faciales consiste en deux tâches successives, tout d'abord d'acquisition de l'image de l'apprenant, via une caméra numérique de type webcam, puis de zoom sur le visage pour éliminer une trop grande quantité d'information inutile pour la reconnaissance par classification, comme l'arrière plan (figure 3).



Figure 3. Extraction des expressions faciales

Reconnaissance des expressions faciales

Pour effectuer cette reconnaissance, nous avons choisi de suivre l'approche connexionniste. Nous pensons tout d'abord que l'expression faciale d'une émotion ne peut pas systématiquement être réduite à un ensemble de caractéristiques géométriques, et qu'une approche plus globale, sur la totalité ou des parties du visage, devrait apporter de meilleurs résultats. En outre, un système connexionniste devrait, suite à un apprentissage de patrons connus et préalablement catégorisés, être en

mesure de généraliser son processus de classification, et ainsi pouvoir traiter avec succès des patrons non connus. Ce module effectue donc une reconnaissance des expressions faciales en classant ainsi chaque image dans une des six catégories d'expressions faciales élémentaires définies par Paul Ekman [2] (nous utilisons également l'expression neutre comme septième catégorie). Deux sous-modules composent ce module de reconnaissance (figure 4). Le sous-module principal est un réseau de neurones artificiels chargé de classer chaque image d'entrée dans l'une des sept catégories d'expressions faciales. Le second module, opérant en amont de l'autre, effectue une réduction de la dimension des images d'entrées. En effet, un réseau de neurones artificiels, pour être capable d'effectuer une généralisation dans sa catégorisation, doit disposer d'entrée dont la dimension est relativement réduite. Une image de visage, même sans arrière plan, présente une dimension beaucoup trop importante (256*256 pixels avec une bonne résolution, soit 65536 données), rendant impossible toute généralisation.

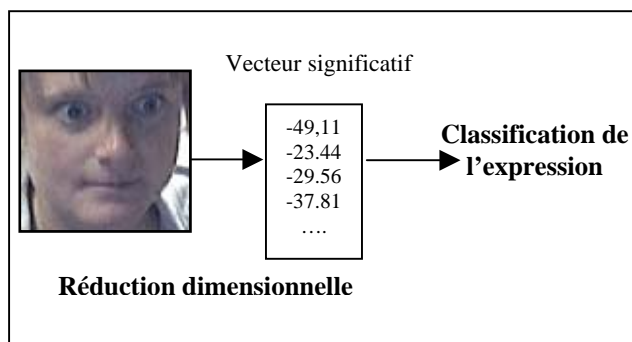


Figure 4. Module de réduction d'image

Module de réduction de données

Parmi les différentes méthodes existantes, [6 ; 7 ; 8], nous avons choisi d'utiliser la décomposition par « Eigenfaces » [7 ; 9 ; 10 ; 11], simple à implémenter et dont l'algorithme est clairement documenté, tout en laissant présager des taux de reconnaissances acceptables (79.3% selon Donato [6]). Cette méthode part du principe qu'il existe une corrélation entre les différents pixels d'une image, ce qui sous-entend l'existence d'une information réduite caractérisant l'intégralité de l'image. En extrayant l'information pertinente, on construit un sous-espace mettant en lumière les variations significatives des images et on réduit les images initiales en effectuant une projection dans ce sous-espace.

Cette méthode construit une nouvelle base à partir de vecteurs propres de la matrice de covariance des images et exprime chaque image comme une combinaison linéaire des vecteurs de cette base. On obtient ainsi une réduction de la taille des vecteurs images, dans la mesure où les coefficients caractérisent l'image (figure 5).

Module de classification des expressions faciales

Une fois l'extraction des informations caractéristiques effectuées, la reconnaissance des expressions faciales est réalisée par un réseau de neurones de type perceptron multicouches.

Son architecture comprend trois couches, à savoir une couche d'entrée, dont la taille correspond à celle des vecteurs réduits des images, une couche cachée, dont il faudra déterminer avec exactitude le nombre de neurones appropriés, et une couche de sortie comportant 7 neurones, et pour chacune des six expressions faciales de bases, et un septième pour l'expression neutre.

Afin de mettre au point l'architecture du réseau, nous avons procédé à plusieurs expérimentations.

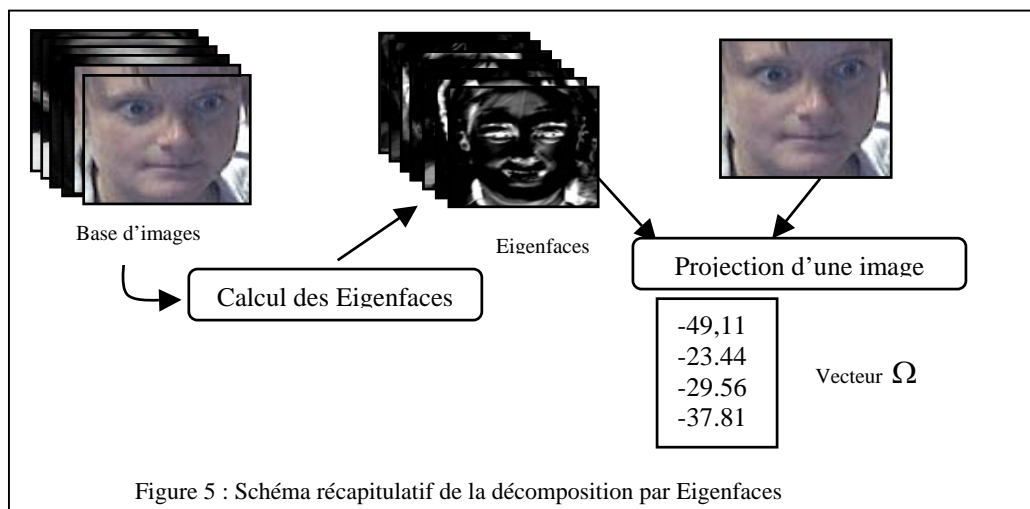


Figure 5 : Schéma récapitulatif de la décomposition par Eigenfaces

Expérimentations et résultats

Objectifs

L'objectif de ces expérimentations est de déterminer l'architecture optimale du réseau de neurones effectuant la classification des expressions faciales. Plus précisément, nous cherchons à trouver le nombre de neurones cachés. Jusqu'à présent, les expérimentations ont été réalisées à partir d'un ensemble d'images statiques, hors de toutes contraintes temporelles. Pour cela, nous avons utilisé la base Cohn-Kanade [12], qui contient près de dix mille images, munies de leur décomposition en actions faciales (FACS). Cela nous permet de réaliser un apprentissage supervisé.

Méthodologie

Pour le moment, nous travaillons seulement avec trois différentes expressions, « neutre », « joie », « surprise ». Nous avons ainsi construit nos eigenfaces à partir d'un ensemble de 97 images (chaque images représentant une des trois expressions) de 32 sujets. De plus, afin d'être sûr de la qualité de décomposition, nous avons fixé le nombre d'eigenfaces à 75, ce qui permet une décomposition de bonne qualité (évaluée via la reconstruction de l'image). Pour ce qui est du réseau, nous avons construit grâce à Matlab un perceptron multicouche utilisant l'algorithme de rétro-propagation résiliente [13], et, en faisant varier le nombre de neurones cachés entre 16 et 30, nous testons chaque architecture. Cette gamme de valeur contient les valeurs empiriques admises dans la littérature [14].

Évaluation de la qualité de l'apprentissage

Afin d'améliorer la qualité de l'apprentissage, et les capacités de généralisation du réseau, nous utilisons un

jeu de validation. Cela signifie que, pour chaque processus d'apprentissage, deux ensembles d'images sont utilisés : le jeu d'apprentissage et le jeu de validation, qui permet d'arrêter l'apprentissage dès lors que le réseau commence à mémoriser le jeu d'apprentissage (ce qui se traduit par l'augmentation du taux d'erreur associé au jeu de validation). Nous utilisons alors le taux d'erreur du jeu de validation (moins bon) comme taux d'erreur global pour cet apprentissage. De plus, nous améliorons cette opération grâce à un processus de validation croisée. En fait, à partir d'un ensemble de patrons d'apprentissage, nous construisons le jeu d'apprentissage (2/3 de l'ensemble des patrons), et le jeu de validation (1/3). La validation croisée est alors réalisée en permutant, après chaque apprentissage, les différents tiers servant à la validation. Le taux d'erreur global de l'architecture est alors calculé comme étant la moyenne des trois taux d'erreurs. (figure 6). Actuellement, nous utilisons un ensemble de patrons d'apprentissage composé de 296 images (105 de joie, 103 de neutres et 88 de surprise), triées aléatoirement.

Résultats

Cette section présente les différents résultats obtenus lors des expérimentations. Nous y montrons le taux global d'erreur et les résultats pour chaque architecture, issus d'un apprentissage avec l'algorithme de rétro-propagation résilient. Afin de simplifier la compréhension des résultats, nous détaillons les calculs effectués pour la première architecture, un réseau composé de 16 neurones cachés (figure 7).

La procédure de validation croisée entraîne trois apprentissages successifs. Le premier d'entre eux (figure 8) aboutit à un taux d'erreur $\mathcal{E}_1 = 0.0401$, qui correspond à l'erreur de validation.

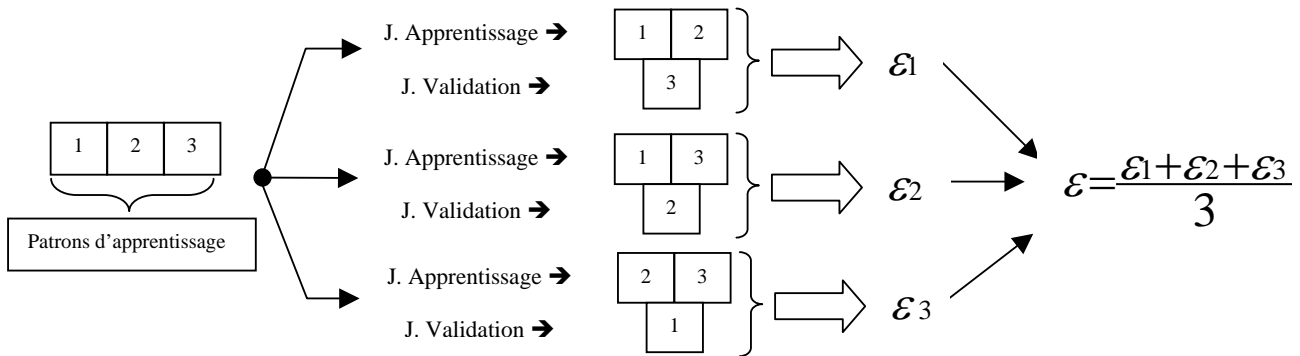


Figure 6 : Approche par validation croisée

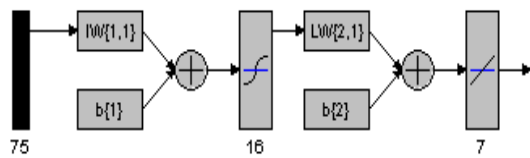


Figure 7 : Architecture du perceptron 75 – 16 - 7

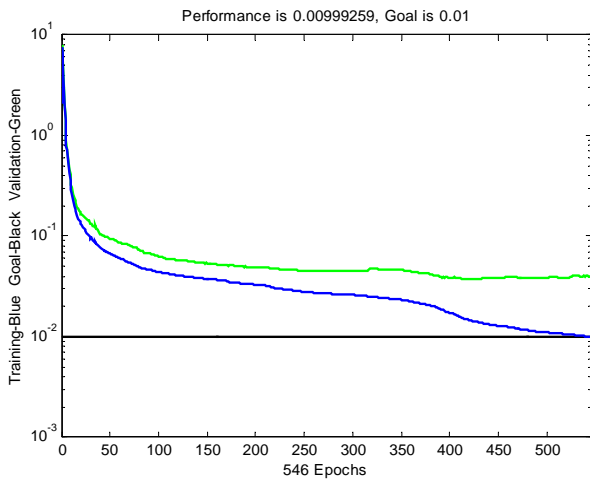


Figure 8 : Courbes du premier apprentissage

Le second apprentissage, ainsi que le troisième, aboutissent à des taux d'erreurs $\mathcal{E}_2 = 0.0158$ et $\mathcal{E}_3 = 0.0135$. Ainsi, nous obtenons un taux d'erreur global, pour 16 neurones cachés, de $\mathcal{E} = 0.0231$.

Afin d'obtenir davantage de résultats, nous avons réalisé plusieurs simulations, en faisant varier le nombre de neurones cachés. Les principaux résultats sont présentés dans la table n°1.

Cette table montre le taux d'erreur résultant du processus de validation croisée, pour chaque architecture, et ce répété 10 fois (seules 3 simulations sont présentées). Nous pouvons voir que la valeur en caractère gras,

0.0231, correspond au résultat de la première simulation (16 neurones cachés). Cette valeur est celle résultante de l'exemple ci-dessus. D'une manière générale, ces résultats montrent que le réseau est effectivement capable d'apprendre correctement ses images, tout en maintenant de bonnes possibilités de généralisation, en évitant de mémoriser les patrons fournis en entrées.

Neurones cachés	Numéro de Simulation			Moyenne
	1	2	10	
16	0,0231	0,0231	0,0202	0,0218
17	0,0261	0,0372	0,0242	0,0267
18	0,0205	0,0222	0,0236	0,0256
19	0,0247	0,0220	0,0182	0,0247
20	0,0178	0,0316	0,0396	0,0288
21	0,0324	0,0258	0,0355	0,0256
22	0,0358	0,0362	0,0239	0,0318
23	0,0234	0,0212	0,0394	0,0292
24	0,0201	0,0307	0,0517	0,0305
25	0,0290	0,0232	0,0273	0,0346
26	0,0329	0,0371	0,0418	0,0327
27	0,0321	0,0550	0,0341	0,0315
28	0,0347	0,0218	0,0514	0,0320
29	0,0290	0,0223	0,0393	0,0325
30	0,0373	0,0276	0,0478	0,0418

Table 1 : Taux d'erreur du réseau

Évaluation de la qualité du processus de généralisation

Outre la capacité d'apprentissage du réseau, nous souhaitons analyser les capacités du réseau en situation d'apprentissage. Nous utilisons pour cela deux tests différents : le premier vise à étudier la capacité de classification du réseau sur images connues (déjà apprises), tandis que le second s'intéresse aux capacités de généralisation du réseau.

Le premier test

Nous utilisons 48 images (14 de joie, 17 neutres et 17 de surprise) déjà apprises. Le test consiste simplement à simuler les reconnaissances de ces images, une fois l'apprentissage terminé, et à comparer l'erreur résultante (entre la sortie attendue et la sortie obtenue). Ces résultats sont dans la table 2.

Neurones cachés	Joie	Neutre	Surprise
16	34%	26%	22%
17	39%	46%	15%
18	30%	44%	41%
19	33%	31%	26%
20	45%	32%	7%
21	17%	42%	30%
22	39%	23%	30%
23	45%	23%	41%
24	29%	47%	19%
25	10%	38%	19%
26	22%	14%	37%
27	18%	29%	44%
28	13%	19%	41%
29	22%	32%	22%
30	25%	31%	44%

Table 2 : Reconnaissance d'images connues

Ces résultats montrent que les réseaux sont effectivement capables de catégoriser des images qu'il a déjà apprises, ce qui est normal. Néanmoins, nous constatons que les résultats obtenus sont moins bons qu'escomptés. En effet, nous aurions pu nous attendre à des taux proches de 100%. La différence provient d'un ensemble d'images qui sont proches d'une frontière entre deux classes, c'est-à-dire d'images très similaires, mais devant être catégorisées différemment. Cela entraîne des erreurs.

Le second test

Le second test a pour objectif d'établir les capacités de généralisation du réseau. Pour cela, nous simulons la reconnaissance d'images non-apprises, représentant de nouveaux sujets, avec une expression connue (joie, surprise, neutre). Nous utilisons pour ce test 70 images (35 joie, 26 neutres, 9 surprise). Les résultats se trouvent dans la table 3.

Comme nous pouvons le voir, ces résultats sont beaucoup plus faibles que les précédents, se situant globalement entre 10 et 45%, dépendamment de l'architecture. Plusieurs éléments peuvent expliquer ces résultats. Tout d'abord, le nombre d'images.

Neurones cachés	Joie	Neutre	Surprise
16	86%	94%	81%
17	81%	100%	85%
18	86%	92%	88%
19	88%	92%	83%
20	88%	100%	71%
21	86%	92%	90%
22	83%	98%	85%
23	88%	100%	85%
24	86%	94%	83%
25	83%	96%	69%
26	95%	96%	90%
27	86%	92%	83%
28	88%	98%	81%
29	88%	96%	88%
30	83%	96%	90%

Table 3 : Reconnaissance d'images connues

En n'utilisant qu'un nombre très réduit d'images servant à l'apprentissage (296) ne peut que difficilement aboutir à une généralisation solide et fiable. De plus, en analysant plus en détails les résultats fournis par le réseau, on constate que les erreurs rencontrées dans le test 1 sont également présentes ici, et que certaines images, possédant un cadrage différent des autres, sont systématiquement mal classifiées. Ce n'est pas surprenant, les réseaux de neurones artificiels étant très sensibles à ces variations.

Conclusion

Plusieurs travaux ont porté sur la reconnaissance des émotionnels à travers les expressions faciales en utilisant une approche connexionniste [15]. Toutefois, ces travaux ont buté sur le nombre d'entrées du réseau de neurones limitant ainsi leur performance surtout lorsque la reconnaissance doit se faire en temps réel. La couche perception de Emilie-2 offre une meilleure performance avec des entrées réduites. De plus, la couche perception est la première de 3 couches importantes de l'agent qui lui permettent de réalisant une boucle pédagogique allant de la reconnaissance d'émotions à travers l'expression faciale jusqu'à la remédiation de ces émotions à passant par une interprétation et un diagnostic judicieux de ces émotions. Le diagnostic permet de mettre en évidence des facteurs cognitifs ayant causés l'émotion permettant ainsi au tuteur de prendre, s'il y a lieu des actions pédagogiques appropriées.

Références

[1] Reynolds, C. and Picard, R. W. 2001. Designing for Affective Interactions. In Proceedings from the 9th International Conference on Human-Computer Interaction, New Orleans.

- [2] Ekman, P. and Friesen, W. V. 1978. *The Facial Action Coding System: A Technique for the Measurement of Facial Movement*. Palo Alto: Consulting Psychologists Press.
- [3] Nkambou, R., Laporte, Y, Yatchou, R. et Gouradères, G. (2002). "Embodied Emotional Agent and Intelligent Training System". In: Abraham, A., Jain, L., & Kacprzyk, J. (Eds), "*Recent Advances in Intelligent Paradigms and Applications*". Chapter 11, pp. 233-253. Springer-Verlag
- [4] Tchetaigni, J. and R. Nkambou. *A framework for diagnosis and epistemological remediation in learning logic programming*. In *International Conference on Computer Aided Learning in Engineering Education*. 2004 <http://www-clips.imag.fr/calie04>.
- [5] Tchetaigni, J., Nkambou R. and Kabanza (2004). *Epistemological Remediation in Intelligent Tutoring Systems*. In: proceedings of the 17th *International Conference on Industrial and Engineering Applications of Artificial Intelligence and Expert System*. Springer-Verlag, Berlin.
- [6] Donato, G. et al. 1999. *Classifying Facial Actions*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 21, no 10. p974-989.
- [7] Turk, M., Pentland, A. 1991. *Eigenfaces for recognition*. Journal of Cognitive NeuroScience 3(1): 7-86.
- [8] Penev, P.S., Atick, J.J. 1996. *Local feature analysis : a general statistical theory for object representation*. Network : Computation in Neural Systems, 7(3) :477-500.
- [9] Li, Zhifeng, Tang, Xiaoou. 2002. *Eigenface recognition using different training data sizes*. InfoSecu'02.
- [10] Calder, A.,J., et al. 2001. *A principal component analysis of facial expression*. Vision Research, Volume 41, Issue 9, April 2001, Pages 1179-1208.
- [11] Heseltine, T., Pear, N., Austin, J. 2002. *Evaluation of image pre-processing techniques for eigenface based face recognition*. Proceedings of the Second International Conference on Image and Graphics, SPIE vol. 4875, pp. 677-685.
- [12] Kanade, T. Cohn, J. F., & Tian, Y. (Mars 2000). *Comprehensive Database for Facial expression Analysis*. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition (FG'00). Grenoble, France.
- [13] Riedmiller, M., and H. Braun, « *A direct adaptive method for faster backpropagation learning: The RPROP algorithm* » Proceedings of the IEEE International Conference on Neural Networks, San Francisco, 1993
- [14] Lepage, R., Solaiman, B. 2003. *Les réseaux de neurones artificiels et leurs applications en imagerie et en vision par ordinateur*. Montréal : Presses de l'École de technologie supérieure. ISBN 2-921145-40-5.
- [15] Lisetti, C. L., Schiano, D. J. 2000. *Automatic Facial Expression Interpretation : Where Human-Computer Interaction, Artificial Intelligence and Cognitive Science Intersect*. Pragmatics and Cognition. Special issue on Facial Information Processing : A multidisciplinary Perspective (8) : 185-235