

Reconnaissance automatique des émotions faciales: vers une application en sciences humaines

Ahmed Fnaiech, Hanene Sahli, Mounir Sayadi

Laboratoire Signal, Image et Maîtrise d'Énergie

École Nationale Supérieure D'Ingénieurs de Tunis - Université de Tunis

Il est évident que le visage humain est le meilleur moyen pour l'identification des personnes, mais il est, aussi, considéré comme un miroir qui reflète les sentiments et les émotions intérieures de la personne et permet ainsi de détecter les besoins de chaque personne et étudier son comportement. Nous pouvons également savoir, à partir du visage humain, les goûts de chaque personne et prévoir ses réactions. En effet, la détection, la classification et la caractérisation du visage sont des secteurs en plein essor. Ce domaine touche l'état psychologique de l'être humain, son comportement et ses réponses.

La reconnaissance automatique des émotions faciales, qui sont dues à des petits mouvements musculaires du visage, permet d'extraire des informations importantes du visage humain pour bien décrire son état cognitif. Ces informations doivent être invariantes et pertinentes même dans le cas de changement d'orientation du visage et aussi de port de lunettes ou présence de barbes.

L'intérêt pratique de la reconnaissance automatique des émotions peut être dans la surveillance à domicile des personnes, surtout les personnes âgées, les enfants ou les handicapées, la détection des abus cybernétiques, le contrôle de sécurité dans les lieux publics

Beaucoup pensent qu'il est difficile de définir le concept d'émotion. La raison en est sans doute que l'émotion n'existe pas, du moins, comme entité psychologique ou comme catégorie aux frontières clairement définies. Tout comme pour le psychologue, la mémoire n'existe pas en tant que telle. Dans les deux cas, le concept recouvre une multitude de phénomènes que le sens commun regroupe dans une large catégorie. Pour l'émotion, ces phénomènes peuvent comprendre une sentimentalité romantique à la lecture d'une histoire d'amour, une crise de colère explosive ou un rougissement lors d'une prise de parole en public. Pour la mémoire, il peut s'agir du souvenir de nos expériences passées (mémoire autobiographique), de notre capacité à apprendre du vocabulaire (mémoire sémantique) ou à maintenir à la conscience une liste de mots (mémoire de travail). D'un point de vue psychologique, ces différents phénomènes, même s'ils sont regroupés par le label commun émotion ou mémoire, sont sous-tendus par des processus très différents. La recherche a montré que cette définition convient parfaitement au phénomène émotion. Les émotions sont en effet constituées d'un ensemble de processus plus ou moins coordonnés et plus ou moins présents selon la nature de l'émotion. Si on définit ces processus et la manière dont ils interagissent, on obtient alors une définition précise et opérationnelle du concept émotion, qui le rend apte à une étude scientifique rigoureuse. Généralement, c'est une situation particulière qui va attribuer une signification émotionnelle. Cette étape est désignée par l'évaluation émotionnelle (Fig. 2). Elle constitue une condition nécessaire à l'émotion. Elle est particulièrement pertinente pour la clinique. En effet, une caractéristique souvent rencontrée chez les personnes souffrant de troubles psychologiques est que des situations ou objets particuliers vont susciter chez elles des réponses émotionnelles importantes

alors que ces mêmes situations ou objets laisseront la plupart des autres personnes indifférentes. Le cas des phobies simples est particulièrement exemplaire à cet égard: un élément anodin, par exemple, un pigeon, provoquant une peur extrême chez le phobique [2].

Quand une signification émotionnelle positive ou négative est activée, l'organisme réagit immédiatement et dans toutes ses composantes pour faire face au défi que constitue le caractère émotionnel de la situation. Il peut s'agir d'une opportunité à saisir, d'un danger à éviter ou à tenir à l'écart, d'un obstacle à éviter. Quoi qu'il en soit, l'organisme humain se prépare à interagir d'une certaine manière avec son environnement: fuir, se cacher, aller vers, aller contre, etc. C'est la tendance à l'action. Pour le phobique, la tendance à l'action est clairement la fuite et l'évitement de l'objet phobogène [2].

En traitement d'images, les méthodes utilisées pour la reconnaissance automatique des émotions faciales sont de natures variées: les réseaux de neurones artificiel (classiques et profonds), les machines à support vectoriel (SVM), les modèles de Markov cachés (HMM), la méthode de la composante principale (ACP).

Fig. 2: Les composantes du processus émotionnel humain

Quelques travaux ont montré que l'utilisation des yeux et de la bouche est suffisante pour la reconnaissance. Plusieurs méthodes de reconnaissance des émotions

utilisent un système d'apprentissage profond à partir de la fusion de la parole et de la vidéo [3][4]. Et d'autres méthodes nécessitent que la personne visualisée soit dans une posture droite parfaitement alignée face à la caméra, ce qui peut être considéré comme un inconvénient. D'où, plusieurs autres travaux se sont intéressés à la détection des émotions dans le visage humain dans les conditions réelles sous différentes orientations: vers la droite, vers la gauche, en haut et en bas.

Parmi les 24 émotions définies par Robert Plutchik [5], uniquement 6 émotions sont généralement étudiées: la joie, la peur, la tristesse, la surprise, la colère et le dégoût, en plus de l'émotion neutre. Les spécificités de chaque émotion sont indiquées dans le tableau suivant [5][9]:

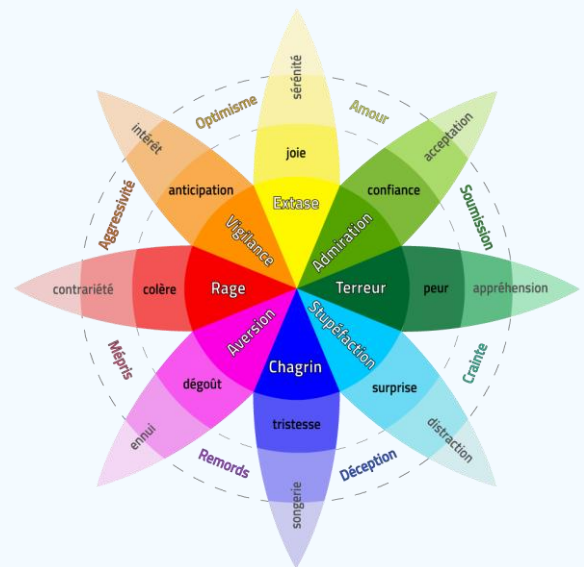
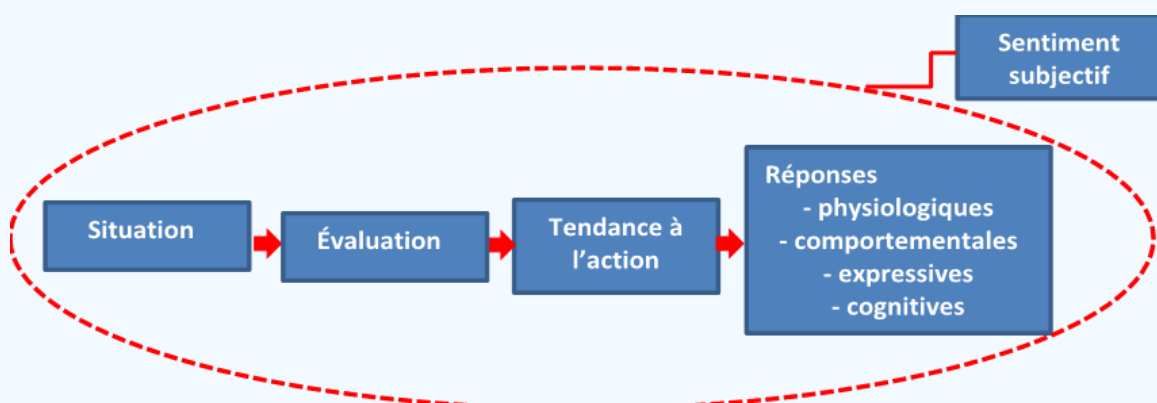


Fig. 1: la roue des émotions de Robert Plutchik [5]



Emotion	Spécifié
La joie 	Facilement identifiables et interprétées en tant que des messages relatifs au plaisir ou en général à une situation positive. Les exemples des expressions de joie sont les émotions les plus courantes en photographie. Elles sont aisément produites à la demande de la personne. En effet, les expressions de joie peuvent cacher des sentiments dans le but de tromper ou manœuvrer les vis-à-vis. Comme le cas des hommes politiques et des célébrités dans le monde des médias.
La tristesse 	Conçues comme l'opposé de celles de la joie. Elles présentent un message lié à la perte, au malaise, à la douleur, à l'abandon, etc. Beaucoup de psychologues indiquent que les gens ayant les visages tristes présentent une intensité plus faible que ceux qui pleurent. Bien que pleurer avec larmes soit un terrain commun concomitant des expressions tristes, les larmes ne sont indicatives d'aucune émotion particulière (les larmes de joie).
La colère 	Souvent enregistrées à cause du stress et des frustrations quotidiennes. C'est le premier sentiment conduisant à une éventuelle agression interpersonnelle. Il porte des messages d'hostilité, d'opposition et d'attaque potentielle. La colère est une réponse commune aux expressions de fureur créant une probabilité de conflit dangereux. L'expression non contrôlée de la colère exerce un effet toxique sur la personne fâchée, la colère chronique semble être associée à certains modes de comportement qui correspondent à des résultats malsains.
Le dégoût 	Les expressions de dégoût sont souvent des réponses corporelles vis-à-vis des situations révoltantes ou écœurantes, comme la chair en décomposition, les matières fécales, les insectes dans les aliments, ou d'autres matières en décomposition qui sont considérées comme incommestibles. Les odeurs désagréables conduisent souvent à des réactions de dégoût.
La peur 	Des expressions de peur ne sont pas souvent enregistrées dans les sociétés sécuritaires. En général, elles sont déclenchées par des situations de violence interpersonnelle ou de danger probable. Elles fournissent des informations sur un danger imminent ou un risque majeur et s'accompagnent d'une émotion négative. Une fois la menace évitée, cette émotion disparaît automatiquement. Le comportement ainsi que les fonctions cognitives sont compromis, pendant que l'évasion devient le but impératif. L'anxiété est liée à la peur et peut impliquer certaines des mêmes réponses corporelles, mais elle est une humeur à plus long terme et ses déclencheurs ne sont pas aussi immédiats.
La surprise 	Les expressions de surprise sont passagères et difficiles à détecter ou enregistrer en temps réel. Elles se produisent presque toujours en réponse aux événements imprévus et elles donnent des messages de situations inattendues ou soudaines. Sachant que l'expression de surprise est de courte durée, elle est souvent suivie d'autres émotions en réponse au sentiment de surprise telles que la joie ou la peur.
Le cas neutre	L'expression faciale neutre se manifeste par l'absence d'émotions affichées sur le visage: les différents muscles du visage sont donc au repos et détendus au maximum

Il est connu que chaque image faciale d'une personne est caractérisée par un ensemble de points spécifiques qui permettent l'identification de son visage. En première phase, 18 points caractérisant parfaitement le visage humain sont détectés, ce qui nous permet de définir 153 distances possibles entre ces points. Ces distances n'ont pas tous un pouvoir de caractérisation des émotions à détecter. Pour cela, un ensemble de distances optimales est choisi en utilisant un critère statistique appelé critère de caractérisation des attributs de Fisher (basé sur le rapport des variances inter-classe par rapport à la variance intra-classe [1]. Généralement, la classification de l'émotion est réalisée en utilisant des réseaux de neurones ayant comme entrées les distances optimisées.

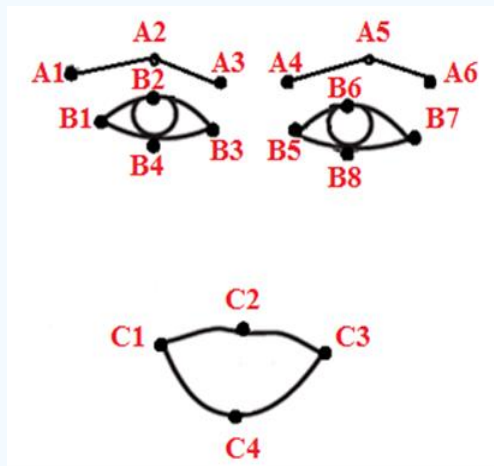


Fig. 3: Les 18 points caractérisant le visage humain

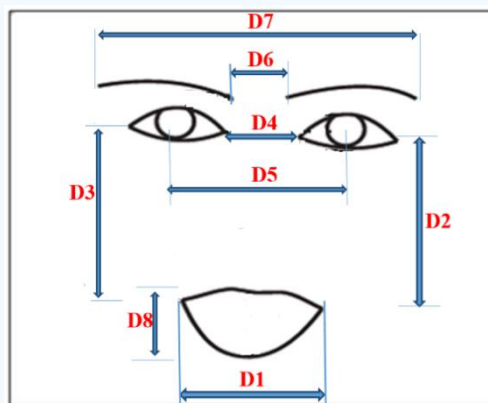


Fig.4: Définition des distances Di les plus utilisées

Dans le processus de la reconnaissance des émotions faciales, plusieurs phases sont nécessaires: une collection de la base de données des différentes émotions est élaborée. Ensuite, une phase de calcul et de calibrage est nécessaire pour marquer les points significatifs sur les visages humains. Une projection 3D- 2D peut être appliquée pour extraire les distances représentatives des émotions traitées. Enfin, un classifieur est utilisé pour discriminer les différences classes des émotions étudiées.

Les résultats réalisés dans le Laboratoire SIME à l'ENSIT, en se basant sur le critère du degré de caractérisation, ont montré que les distances les plus significatives pour la reconnaissance automatique des émotions considérées sont celles représentées dans la figure suivante et surtout D1 (Distance entre les coins gauche et droit de la bouche), D6 (Distance entre les deux coins intérieurs des sourcils) et D8 (Distance entre les coins haut et bas de la bouche) [1][8]. Un apprentissage a été fait avec 50 % des données collectées à partir de plusieurs images faciales de plusieurs personnes. Un réseau de neurones artificielles multicouches a été entraîné avec les 3 distances optimales choisies. Ensuite, pour exploiter la caractéristique de généralisation et de prédiction du réseau de neurones, un test a été réalisé avec le reste des 50 % des données que le réseau de neurones ne connaît pas. La matrice de confusion des résultats de classification des 7 émotions considérées est présentée dans la Table 1. On remarque quelques erreurs de confusion de plus de 3 % spécialement entre la tristesse et la colère, la peur et le dégoût, le dégoût et la tristesse et la joie et la surprise.

Un taux moyen de bonne reconnaissance de 88.43 % a été atteint. Ce taux diminue légèrement en cas de changement de l'orientation du visage.

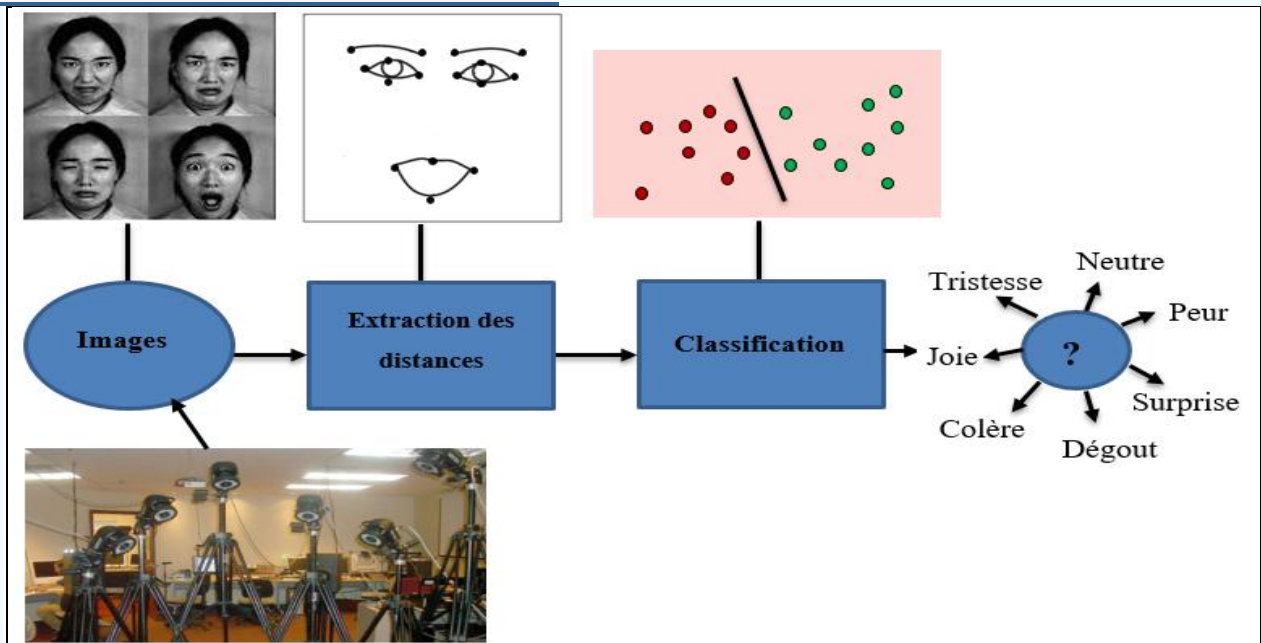


Fig. 5: Schéma synoptique du système réalisé de reconnaissance automatique des émotions faciales

		Prédiction						
Emotion réelle		Colère	Joie	Dégoût	Surprise	Peur	Tristesse	Neutre
	Colère	87,61%	2,54%	2,64%	2,00%	1,68%	2,25%	1,29%
	Joie	1,36%	88,46%	2,57%	3,18%	2,36%	1,75%	0,32%
	Dégoût	2,86%	2,68%	82,93%	3,04%	3,46%	3,29%	1,75%
	Surprise	1,50%	1,64%	2,11%	92,89%	0,89%	0,96%	0,00%
	Peur	2,11%	1,36%	9,96%	0,18%	83,64%	2,75%	0,00%
	Tristesse	3,36%	1,61%	2,82%	1,00%	2,79%	87,39%	1,04%
	Neutre	1,18%	0,43%	1,14%	0,04%	0,07%	1,00%	96,14%

Table 1: Matrice de confusion de résultats de classification des émotions considérées

Bibliographie

- [1] Ahmed Fnaiech; Sahli, Hanene; Sayadi, Mounir; Gorce, Philippe. 2021. "Fear Facial Emotion Recognition Based on Angular Deviation" Electronics vol. 10, no. 3, February 2021.
- [2] P. Philippot, "Emotion et psychothérapie", Edition Wavre, Belgium: Mardaga, 2011.
- [3] Hossain, M.S.; Muhammad, G. *Emotion recognition using deep learning approach from audio-visual emotional big data*. Inf. Fusion 2019, 49, 69–78.
- [4] Kaya, H.; Gürpınar, F.; Salah, A.A. *Video-based emotion recognition in the wild using deep transfer learning and score fusion*. Image Vis. Comput. 2017, 65, 66–75.
- [5] Donaldson, Melissa. "Plutchik's wheel of emotions, 2017.
- [6] Nguyen, D.H.; Kim, S.; Lee, G.S.; Yang, H.J.; Na, I.S.; Kim, S.H. *Facial Expression Recognition Using a Temporal Ensemble of Multi-level Convolutional Neural Networks*. IEEE Trans. Affect. Comput. 2019, 33.
- [7] Giannopoulos, Panagiotis, Perikos, Isidoros, et Hatzilygeroudis, Ioannis. *Deep learning approaches for facial emotion recognition: A case study on FER-2013*. In: Advances in hybridization of intelligent methods. Springer, Cham, 2018. p. 1-16.
- [8] Ahmed Fnaiech; « Optimisation de la détection des émotions humaines dans les images faciales 3D », thèse de Doctorat, soutenue le 21 décembre 2021, ENSIT, Université de Tunis.
- [9] Ekman, Paul, Friesen, Wallace V., Hager, J. *Investigator's Guide to the Facial Action Coding System (FACS)*. Salt Lake City, UT: A Human Face, 2002.

Biographie



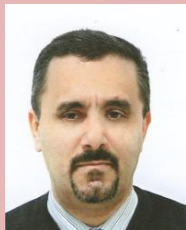
Ahmed Fnaiech

Docteur en Traitement du signal et image de l'ENSIT, Université de Tunis et membre du Laboratoire Signal Image et Maîtrise de l'Energie à l'ENSIT.



Hanene Sahli

Maître-Assistant en Traitement du signal et image à ISSAT, Université de Kairouan et membre du Laboratoire Signal Image et Maîtrise de l'Energie à l'ENSIT. Ses travaux de recherche sont concentrés autour du traitement des signaux et des images biophysiques



Mounir Sayadi

Professeur en Traitement du signal et image de l'ENSIT depuis 2015, et Directeur du Laboratoire Signal Image et Maitrise de l'Energie depuis 2018. Ses travaux de recherche du traitement du signal et des images médicales et aériennes et de la digitalisation et modélisation des systèmes