

M

on ordinateur sait quand je souris!

Un ordinateur peut-il reconnaître les expressions faciales d'une personne ?

Oui, sous certaines conditions, la technologie le permet. Cependant, beaucoup de recherches sont encore nécessaires pour gérer toutes les subtilités d'un visage expressif. Voici pourquoi...

Par France Laliberté, Ph.D
et Langis Gagnon, Ph.D
Équipe Vision et imagerie, CRIM

Expression ou émotion ?

Tout d'abord, il est important de faire la distinction entre la reconnaissance d'expressions faciales et la reconnaissance d'émotions. Les émotions résultent de plusieurs facteurs et peuvent être révélées par la voix, la posture, les gestes, la direction du regard et les expressions faciales. Par contre, les émotions ne sont pas la seule origine des expressions faciales. En effet, celles-ci peuvent provenir de l'état d'esprit (par ex. la réflexion), de l'activité physiologique (comme la douleur ou la fatigue) et de la communication non verbale (notamment émotion simulée, clin d'œil, froncement des sourcils). Néanmoins, sept émotions de base correspondent à une expression faciale unique, quelles que soient l'ethnicité et la culture du sujet : colère, dégoût, étonnement, joie, mépris, peur et tristesse¹. La reconnaissance

des expressions faciales consiste à classer les déformations des structures faciales et les mouvements faciaux uniquement à partir d'informations visuelles. La reconnaissance des émotions, quant à elle, est une tentative d'interprétation qui requiert une information contextuelle plus complète.

Mesure des expressions faciales

Les expressions faciales consistent en une déformation temporaire (entre 0.06 et 5 secondes) des structures faciales. Il existe deux types de structures faciales : permanentes (par ex. yeux, sourcils, bouche, poils, rides profondes) et transitoires (rides, gonflement). Les structures faciales permanentes sont toujours présentes sur le visage, mais peuvent être déformées par les expressions faciales. Les structures faciales transitoires, elles, sont créées par les expressions



L'image de gauche représente un sourire sincère, qui est composé de l'unité d'action 6 (provoqué par la contraction d'un muscle autour des yeux). On l'appelle le sourire de Duchenne, en l'honneur d'un physiologiste et neurologue du 19^e siècle : Guillaume Duchenne. L'image de droite représente un sourire social, appelé aussi « professionnel » ou « PanAmerican », en l'honneur des agents de bord de la défunte compagnie aérienne du même nom.

faciales; elles sont principalement localisées sur le front et autour de la bouche et des yeux.

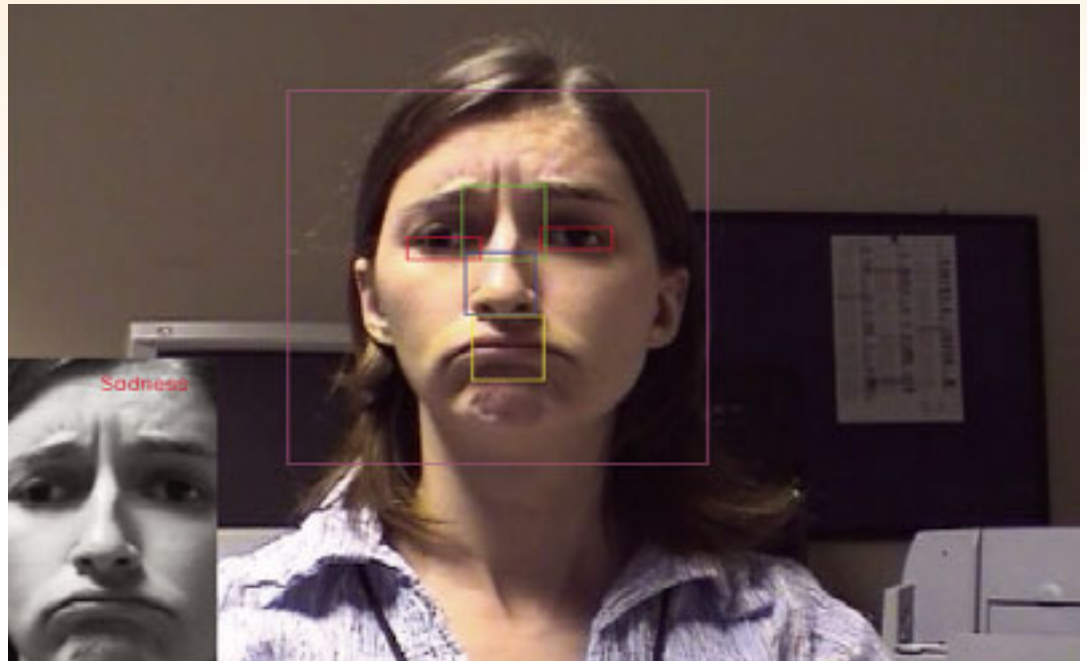
La terminologie utilisée pour décrire les expressions faciales est basée sur la position, l'intensité et la dynamique (synchronisation et évolution) des actions faciales (déformations et mouvements faciaux). L'intensité des actions faciales peut être mesurée en déterminant la déformation géométrique des structures faciales ou la densité des rides apparaissant dans certaines régions du visage. La dynamique des actions faciales contient une information significative et il est donc préférable de mesurer les expressions faciales à partir d'une vidéo au lieu d'une seule image.

Analyse des expressions faciales

Un système d'analyse automatique des expressions faciales s'effectue généralement selon les étapes suivantes : normalisation et segmentation du visage, extraction et représentation des structures faciales et enfin, classification (reconnaissance).

La normalisation sert à atténuer les changements d'apparence du visage qui sont causés par la position de la tête ou les variations d'éclairage. La segmentation du visage permet d'isoler les structures faciales ou d'effectuer une séparation entre les visages et le fond de la scène.

Il existe plusieurs méthodes d'extraction des structures faciales que l'on peut regrouper en deux grandes classes : les méthodes basées sur l'extraction des déformations du visage et celles basées sur l'extraction du mouvement. Les méthodes d'extraction des déformations traitent chaque image d'une séquence vidéo individuellement et ne peuvent donc pas reconstruire le mouvement des pixels. Les méthodes d'extraction du mouvement mettent



Interface du système développé au CRIM pour la reconnaissance d'expressions faciales et de visages. Les structures faciales détectées et l'expression reconnue sont affichées.

l'accent sur les changements faciaux causés par les expressions faciales et ne peuvent donc pas être appliquées aux images individuelles. Il est plus difficile d'analyser une expression faciale à partir d'une seule image qu'à partir d'une séquence d'images.

On constate une différence entre la reconnaissance et l'interprétation des expressions faciales. En effet, l'interprétation consiste à classer les expressions faciales dans une des sept émotions de base mentionnées plus haut, alors que la reconnaissance consiste à classer les actions faciales, selon leur position et leur intensité, en classes visuelles. Cette étape peut être suivie ou non d'une interprétation selon les émotions de base. Le FACS (*Facial Action Coding System*) est une référence pour la description des actions faciales². Ce système utilise 44 unités d'action pour décrire les actions faciales selon leur position et leur intensité (échelle de 5 degrés de magnitude).

Une expression peut être modélisée à l'aide d'une ou de plusieurs unités d'action. La plupart des travaux utilisent l'approche qui consiste à interpréter directement les expressions faciales selon les émotions de base, car elle est plus simple, même si elle comporte certains désavantages.

Les systèmes actuels

Les systèmes actuels d'analyse d'expressions faciales les plus performants sont ceux qui utilisent plusieurs méthodes complémentaires³. Ils offrent une meilleure performance, car les algorithmes d'extraction mettent l'accent sur des structures faciales différentes et/ou la combinaison des étapes d'extraction, de reconnaissance et d'interprétation produit des erreurs différentes. Les systèmes actuels sont difficilement utilisables dans des conditions réelles d'acquisition des images. De telles conditions incluent des visages légèrement de profil,

à des distances différentes de la caméra, des mouvements importants de la tête ainsi que des variations d'illumination.

De plus, il ne doit y avoir aucune intervention de l'utilisateur et le tout doit se dérouler en temps réel. L'amélioration des systèmes actuels devra se faire principalement au niveau de l'extraction des structures faciales. Quelles que soient les conditions d'acquisition, l'obtention de structures faciales bien localisées, desquelles on devrait obtenir l'intensité et la dynamique des actions faciales, est la clé des futurs systèmes performants d'analyse automatique des expressions faciales. De plus, la conception d'un système fonctionnant en temps réel est essentielle pour des applications d'interaction personne-machine⁴.



Chaque rangée d'images représente la transition entre deux expressions faciales prototypiques. De haut en bas : joie/étonnement, étonnement/peur, peur/tristesse, tristesse/dégoût, dégoût/colère, colère/joie. L'être humain reconnaît les émotions des deux premières colonnes, qui ont une composition de 90 % - 10 % et 70 % - 30 %. Par exemple, pour la première image de la première rangée, l'expression se compose à 90 % de joie et 10 % d'étonnement.

©MRC Cognition and Brain Science Unit.

Applications technologiques possibles

La recherche sur les expressions faciales et les systèmes automatiques qui en découlent ont une multitude d'applications technologiques possibles. Tout d'abord, la mesure des expressions faciales est utilisée en recherche comportementale, car elle indique l'activité émotionnelle de manière moins intrusive et moins coûteuse que l'électroencéphalogramme, l'électromyographie, les systèmes de neuromodulation avancés et l'imagerie cérébrale.

Deuxièmement, dans un contexte d'enseignement à distance où l'enseignant ne voit pas ses étudiants, il est important de lui fournir un écho de leur réaction sous forme d'information de haut niveau (par ex. une moyenne des expressions faciales de la classe). Troisièmement, toutes les interfaces personne-machine (jeux, bureautique, logiciels d'apprentissage) bénéficieraient d'une reconnaissance des expressions faciales pour adapter le contenu (comme le niveau de difficulté) à la réaction du sujet en temps réel. Quatrièmement, les connaissances acquises sur les expressions faciales sont récupérées pour la création d'avatars réalistes. Enfin, la détection des expressions faciales pourra être utilisée dans les voitures, sous forme de dispositif qui avertit le conducteur en cas de perte de vigilance, et dans les avions, comme information supplémentaire pour la boîte noire en cas d'accident.

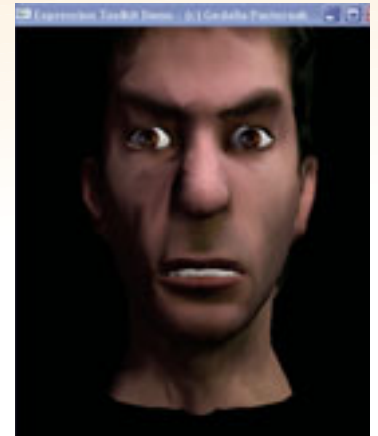
Au CRIM

Des travaux de recherche et développement en reconnaissance et interprétation des expressions faciales à des fins de surveillance vidéo et d'interaction personne-machine sont actuellement en cours au **CRIM**. Ces projets sont menés par l'équipe Vision et imagerie, en partenariat avec le Laboratoire de vision et systèmes numériques de **l'Université Laval**.

Il s'agit des projets MONNET (*MONitoring of extended premises : tracking pedestrians using a NETWORK of loosely coupled cameras*) et COBVIS-D (*a simulation environment for the evaluation of Cephalo-Ocular Behaviour and Visual Search patterns on Drivers*).

La recherche sur les expressions faciales permet de mettre au point des avatars réalistes capables d'afficher des émotions, comme en témoigne cette image tirée d'un système d'animation faciale dans le domaine du logiciel libre.

© The Expression Toolkit



1. Ekman, P., Friesen, W. V., « Constants across cultures in the face and emotion », *Journal of Personality Social Psychology*, vol. 17, n° 2, pp. 124-129, 1971.
2. Ekman, P., Friesen, W. V., « Facial action coding system: a technique for the measurement of facial movement », *Consulting Psychologists Press*, Palo Alto, 1978.
3. Fasel, B., Luetten, J., *Pattern Recognition*, vol. 36, pp. 259-275, 2003.
4. Pantic, M., Rothkrantz, L. J. M., « Automatic analysis of facial expressions: The state of the art », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, n° 12, pp. 1424-1445, 2000.

Cet espace pourrait être le vôtre

pour annoncer vos produits et services auprès du plus important réseau d'experts en TI au Québec.

Appelez-nous en composant le

(514)840-1233