

les changements de couleur, mais aussi les mouvements imperceptibles.

Dans des travaux antérieurs avec d'autres collègues, nous nous étions intéressés à l'amplification de mouvements infimes dans des vidéos. Nous avons utilisé des logiciels spécialisés qui modélisent et calculent le vecteur déplacement de chacun des pixels de l'image. Ce traitement était lourd et sujet à des erreurs. Nous étions maintenant stupéfaits : notre nouvelle approche parvenait à faire la même chose avec des calculs simples et sans recourir aux laborieux vecteurs du mouvement.

De la couleur au mouvement

Pourquoi les changements de couleur accentués grossissaient-ils aussi ces petits mouvements ? Pour le comprendre, nous avons étudié comment le mouvement dans une vidéo se traduit en termes de changements locaux de couleur. Imaginez une balle éclairée par la droite : le côté droit de la balle est lumineux et le côté gauche est sombre. Considérez un pixel donné sur l'écran : si la balle traverse l'écran de gauche à droite, le pixel représente d'abord un point éclairé de la balle, puis s'assombrit à mesure qu'il représente un point de plus en plus à gauche de la balle. La variation dépend de la vitesse de la balle et de la transition abrupte de luminosité entre les côtés gauche et droite de la balle, ce qu'on nomme le gradient de couleur. Mathématiquement, le changement de couleur d'un pixel par unité de temps est égal au produit de la vitesse de l'objet par le gradient de couleur.

Notre algorithme, bien sûr, ne connaît ni les gradients de couleur ni la vitesse des objets dans les vidéos. Toutefois, parce qu'il amplifie le changement de couleur en chaque point quand la balle se déplace d'une fraction de centimètre vers la droite, il amplifie aussi ce mouvement infime de la balle pour le rendre visible à vos yeux. De la même façon, les couleurs des pixels représentant des points particuliers de la poitrine d'un bébé changent à mesure que celui-ci respire, et en accentuant les changements de couleur, on rend beaucoup plus visible le très léger mouvement de la poitrine.

La différence entre nos précédents travaux portant sur les vecteurs déplacement et notre nouvelle approche fondée sur les variations temporelles de couleur est une question de perspective. En faisant une

■ LES AUTEURS



Frédo DURAND est professeur de génie électrique et d'informatique au MIT (Institut de technologie du Massachusetts). Il est spécialiste des techniques numériques de photographie.



William FREEMAN est professeur de génie électrique et d'informatique au MIT. Il étudie l'application de l'intelligence artificielle à la vision par ordinateur.



Michael RUBINSTEIN est chercheur chez Google ; il travaille sur la vision par ordinateur.

L'ESSENTIEL

■ **Même quand ils semblent à nos yeux parfaitement immobiles, les objets et les personnes sont animés de minuscules mouvements.**

■ **En amplifiant les changements de couleur des pixels d'une vidéo image après image, les chercheurs ont créé un « microscope du mouvement » qui rend ces transformations visibles.**

■ **Cet outil peut rendre de nombreux services : suivi du rythme cardiaque ou de la respiration, recherche de défauts de fonctionnement dans des machines, etc.**

analogie avec une rivière, la différence revient à suivre le mouvement de l'eau ou à rester immobile au milieu du courant. C'est ce changement de point de vue qui rend nos calculs beaucoup plus simples. En mécanique des fluides, les physiciens utilisent ces deux approches pour modéliser les écoulements : l'approche dite lagrangienne consiste à suivre une quantité donnée de fluide lors de son déplacement dans l'espace, comme un observateur sur un bateau suit le flot d'une rivière ; à l'inverse, l'approche dite eulérienne fixe un lieu de l'espace et étudie le fluide qui y passe, comme si l'observateur se tenait sur un pont.

Nous avons adopté le point de vue lagrangien dans nos premiers travaux, et nous étions comme des observateurs sur un bateau : les pixels sont suivis sur la vidéo et déplacés de point en point suivant des vecteurs amplifiés. En revanche, notre nouvelle approche considère le changement de couleur en un lieu fixe, comme lorsque l'observateur reste sur le pont. Cette perspective locale ne s'applique qu'à de petits mouvements, mais elle est beaucoup plus simple et robuste. À l'aide de cette technique, un ordinateur traite rapidement une vidéo, alors que nos travaux précédents nécessitaient un long temps de calcul, avec des erreurs fréquentes.

Amplification eulérienne

En 2012, dans un article traitant de cette méthode que nous nommons « amplification vidéo eulérienne », nous avons décrit plusieurs exemples d'utilisation, dont le changement de couleur d'un visage sous l'effet de la circulation sanguine, mais aussi l'observation des mouvements respiratoires d'un nourrisson, amplifiés de telle sorte que les parents d'un nouveau-né pourraient surveiller un signal vidéo accentué afin de vérifier que le bébé bouge.

Nous avons aussi pris une vidéo à grande vitesse d'une guitare dont toutes les cordes vibraient et nous avons sélectionné d'étroites bandes de fréquences autour de certaines notes pour isoler et amplifier une corde à la fois, par exemple l'intervalle entre 72 et 92 hertz pour la corde de mi grave vibrant à 82 hertz. Les autres cordes paraissaient absolument immobiles.

Nous avons créé un site Internet où les gens pouvaient télécharger leurs vidéos et appliquer notre outil d'amplification du mouvement (*pour des exemples, voir <https://lambda.grilab.com/site/>*). Les visiteurs

Ils ont utilisé pour en faire des choses auxquelles nous n'avions pas pensé, ce qui était passionnant. Une personne a posté une vidéo montrant des mouvements fœtaux en fin de grossesse. Une étudiante en art a réalisé une vidéo mettant en évidence les expressions et mouvements imperceptibles de ses amis qui essayaient de rester immobiles.

Notre approche eulérienne a tout de même ses limites. Si un pixel d'entrée donné s'assombrit beaucoup d'une image à la suivante, l'ordinateur accentue trop ce changement, produisant un pixel complètement noir par un effet d'amplification incontrôlé. Ce type de problème crée des halos sombres ou lumineux autour des zones de mouvement. Par ailleurs, les variations de couleur d'entrée liées au bruit du capteur posent aussi problème parce que, même si nous les lisons en

faisant une moyenne locale sur un grand nombre de pixels, le bruit reste amplifié.

Ces résultats nous ont incités, avec notre étudiant en thèse Neal Wadhwa, à développer un nouvel algorithme qui préserve les bénéfices des approches eulériennes simples tout en donnant une meilleure vision quand les changements deviennent extrêmes.

Nous nous sommes rendu compte que l'origine des limites de notre méthode initiale est qu'elle traite de la même façon les bords et les zones uniformes qui les entourent, alors que les bords correspondent à des différences de couleurs beaucoup plus grandes (des gradients plus importants). Ainsi, si vous essayez d'amplifier tous les pixels avec le même facteur et en même temps, vous obtenez des distorsions (des effets de saturation) qui ne correspondent pas à des mouvements réels.

Au lieu d'amplifier avec un même facteur, nous avons opté en 2013 pour une approche plus fine qui utilise des techniques modernes de traitement du signal et de l'image. De façon analogue à une décomposition classique de Fourier où le signal est représenté par une somme de fonctions sinusoïdales, nous modélisons l'image (l'ensemble de ses pixels) par une somme de fonctions appropriées, selon une technique dite de pyramides qui analyse plusieurs échelles de résolution. De la même façon qu'un changement de phase des fonctions sinusoïdales dans une analyse de Fourier correspond à une translation, un changement de phase des fonctions utilisées dans notre approche correspond à un mouvement local dans un groupe de pixels de l'image. Cette technique permet de caractériser plus facilement l'ampleur du mouvement qui sépare deux images

TRANSFORMER LES CHANGEMENTS DE COULEUR EN MOUVEMENT

Sur une vidéo, chaque pixel représente un point d'un objet, une feuille d'arbre par exemple. La couleur de ce pixel change au cours du temps avec le mouvement de la feuille, même de façon imperceptible, parce que la feuille bouge par rapport à la lumière qui l'atteint. Un programme informatique qui amplifie les variations de couleurs d'une image à la suivante amplifie également les minuscules mouvements qui deviennent alors visibles à l'œil nu. Les chercheurs peuvent isoler et amplifier certaines fréquences en jeu, par exemple celles correspondant au tremblement des feuilles ; dans ce cas, on fait ressortir le mouvement des feuilles par rapport à celui de l'ensemble de l'arbre, les branches bougeant à un rythme plus lent.

