



HAL
open science

Cartes de Saillance Spatio-Temporelle basées Contrastes de Couleur et Mouvement Relatif

Olivier Brouard, Vincent Ricordel, Dominique Barba

► **To cite this version:**

Olivier Brouard, Vincent Ricordel, Dominique Barba. Cartes de Saillance Spatio-Temporelle basées Contrastes de Couleur et Mouvement Relatif. Compression et representation des signaux audiovisuels, CORESA 2009, Mar 2009, Toulouse, France. 6 p. hal-00364867

HAL Id: hal-00364867

<https://hal.science/hal-00364867v1>

Submitted on 27 Feb 2009

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.

Cartes de Saillance Spatio-Temporelle basées Contrastes de Couleur et Mouvement Relatif

O. Brouard¹

V. Ricordel¹

D. Barba¹

¹Université de Nantes – Laboratoire IRCCyN – Équipe IVC

Polytech⁷ Nantes, rue Christian Pauc, 44306 Nantes, France

{olivier.brouard, vincent.ricordel, dominique.barba}@univ-nantes.fr

Résumé

Dans cet article, nous proposons une modélisation de saillance spatio-temporelle. La saillance spatiale est obtenue à partir des couleurs et plus particulièrement de leur contraste. La saillance temporelle est évaluée à partir du contraste de mouvement. En effet, une zone temporellement saillante est une zone en contraste de mouvement. Dans l'approche proposée, le contraste de mouvement est déterminé à partir d'une estimation locale et d'une estimation globale du mouvement ; la différence, appelée mouvement relatif, indique les zones temporellement saillantes. La saillance finale est obtenue en combinant les saillances spatiale et temporelle.

Mots clefs

Attention visuelle, contrastes, mouvement relatif, saillance spatio-temporelle, codage vidéo.

1 Introduction

Pour faire face à l'énorme quantité d'informations de notre environnement, le système visuel humain (SVH) possède la faculté de sélectionner une information pertinente localisée spatialement dans le champ visuel parmi toutes celles qui lui parviennent : on parle d'attention visuelle. Du fait de la grande complexité des interactions et des inter-dépendances existants entre les mécanismes du SVH, modéliser l'attention visuelle dans son ensemble reste trop complexe. Une voie réaliste est de modéliser l'attention visuelle pré-attentive. Le modèle proposé doit être capable de déterminer les zones visuellement importantes d'une image et dans notre cas d'une séquence vidéo.

De nombreux facteurs influençant l'attention visuelle ont été identifiés [1] et sont regroupés en deux catégories. La première regroupe toutes les informations spatiales [2] dont les plus susceptibles de stimuler l'attention visuelle regroupent la couleur, l'orientation et la taille. La deuxième catégorie concerne les informations temporelles [3]. Une séquence vidéo contient ces deux types d'informations susceptibles de stimuler l'attention visuelle. C'est pourquoi notre modèle d'attention visuelle pré-attentive doit prendre en compte. Celui-ci se décompose en deux parties,

l'une modélise l'attention visuelle à partir des informations spatiales et l'autre à partir des données temporelles. La dernière étape combine ces deux parties afin d'obtenir une carte de saillance spatio-temporelle. Dans la partie suivante, nous présentons notre méthode de calcul de cartes de saillance spatiale. Dans la partie 3, nous décrivons la méthode utilisée pour obtenir les cartes de saillance temporelle. La combinaison des cartes de saillance spatiale et temporelle est décrite dans la partie 4. Finalement, nous montrons les résultats obtenus dans la partie 5 et concluons.

2 Saillance spatiale basée sur le contraste de couleur

Des informations importantes peuvent être trouvées dans la littérature sur la théorie des couleurs et plus particulièrement sur les attributs des couleurs qui contribuent à rendre un objet visuellement saillant ou non [4,5]. En terme de saillance de couleur, d'autres méthodes d'attention visuelle artificielle se sont concentrées uniquement sur des attributs de couleurs qui ont été signalés en psychologie. Les artistes utilisent ces aspects pour créer des effets de contraste, des mises en avant visuelles et de mobilité dans leurs illustrations. Dans leurs travaux Aziz et Mertsching [2] combinent ces concepts et forment un ensemble de points possibles à mettre en œuvre. Il reste ensuite à décider quelles couleurs vont être saillantes en présence d'un contraste. Les différents points avec mention de la couleur saillante sont énumérés ci-dessous :

1. *Contraste de Saturation* : Un contraste est produit par des couleurs faiblement et fortement saturées. La valeur du contraste est directement proportionnelle à la magnitude de la différence de saturation. Les couleurs fortement saturées tendent à attirer l'attention dans de telles situations, à moins qu'une région faiblement saturée soit entourée par une région fortement saturée.
2. *Contraste d'Intensité* : Un contraste sera visible lorsque des couleurs sombres et lumineuses coexistent. Plus la différence d'intensité est importante, plus l'effet de contraste augmente. Les couleurs lumineuses attirent l'attention dans cette situation, à moins

que la région sombre ne soit entourée par une région lumineuse.

3. *Contraste de Teinte* : La différence des angles de teinte sur le disque des couleurs (cf figure 1) contribue à la création d'un contraste. Une différence importante va manifestement produire un contraste fort. Du fait de la nature circulaire de la teinte, la plus grande différence entre deux valeurs de teinte est de 180° .
4. *Contraste d'Opposants* : Les couleurs situées sur les côtés opposés du disque de teinte produisent une importante valeur de contraste. Cela signifie naturellement que la différence des angles des valeurs de teinte est dans ce cas proche de 180° . Les couleurs situées dans la première moitié du disque de teinte, connues comme la gamme de couleur active, domineront sur le reste des couleurs passives.
5. *Contraste des couleurs Chaudes et Froides* : Les couleurs chaudes, c'est-à-dire rouge, jaune et orange sont visuellement plus saillantes. Ces couleurs sont situées dans les premiers 45° du disque de teinte. Les couleurs chaudes et froides créent un contraste pour lequel les couleurs chaudes restent dominantes.
6. *Dominance des Couleurs Chaudes* : Les couleurs chaudes dominent leur environnement, même si un contraste existe dans cet environnement.
7. *Dominance de la Luminosité et de la Saturation* : Les couleurs fortement lumineuses et saturées sont considérées comme étant attractives sans tenir compte de leurs valeurs de teinte. De telles couleurs ont plus de chances d'attirer l'attention.

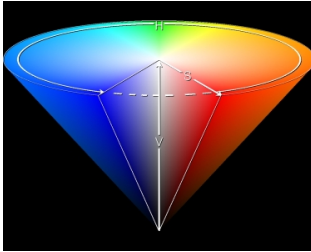


FIG. 1 – Représentation conique de l'espace TSV (HSV).

L'effet de contraste est contrôlé par la valeur de saturation des deux couleurs impliquées dans les situations mentionnées aux points 2 à 5. Les couleurs fortement saturées impliquent des contrastes importants. Notre modèle de saillance des couleurs, basé sur les travaux de Aziz et Mertsching [2], combine tous les points mentionnés ci-dessus. Nous divisons cette procédure en sept étapes, chacune contribuant à la valeur de saillance d'un site s . Les valeurs des différentes composantes de couleur utilisées dans nos calculs pour un site s , sont les valeurs moyennes du bloc situé au centre du tube considéré [7].

2.1 Calcul de la saillance spatiale

Les cinq premières étapes de l'algorithme utilisent un ou les deux facteurs de saturation f_{ij}^{sat} et d'intensité f_{ij}^{int} dans leurs calculs. Les indices i et j représentent respectivement la position du site courant et d'un site voisin (voisinage 8-connexte). La première partie du facteur de saturation f_{ij}^{sat} est obtenue en calculant la moyenne des valeurs de saturation entre le site s_i et le site s_j , l'effet de ce facteur est donc plus important lorsque les deux blocs ont une valeur élevée de saturation et vice versa. La deuxième partie dépend seulement de la saturation du site s_i et détermine une valeur minimale égale à k_{min} , afin de ne pas supprimer l'interaction des blocs avec une saturation proche de zéro. Le reste de la seconde partie est obtenue à partir de la saturation du site s_i et est pondéré par $(1 - k_{min})$. Le facteur pour l'intensité est calculé de la même façon en utilisant la valeur de l'intensité de la couleur du bloc et non la valeur de la saturation. Soient $S(s_i)$ et $I(s_i)$, respectivement les valeurs de saturation et d'intensité du site s , la valeur maximale pour la saturation et l'intensité étant égale à 1, les deux facteurs de saturation et d'intensité sont définis par :

$$f_{ij}^{sat} = \frac{S(s_i) + S(s_j)}{2} \times (k_{min} + (1 - k_{min}) \cdot S(s_i)), \quad (1)$$

$$f_{ij}^{int} = \frac{I(s_i) + I(s_j)}{2} \times (k_{min} + (1 - k_{min}) \cdot I(s_i)), \quad (2)$$

où $k_{min} = 0,21$.

La contribution de la première étape en terme de saillance pour un site s_i est obtenue à partir des deux facteurs de saturation et d'intensité :

$$X_1(s_i) = \sum_{j=1}^{j=p_i} f_{ij}^{sat} \cdot f_{ij}^{int} \quad \forall s_j \in \eta_i, \quad (3)$$

où p_i est la taille du voisinage (8-connexte) et η_i représente l'ensemble des sites voisins de s_i .

La seconde étape collecte les contributions des sites qui ont une valeur de teinte éloignée de celle du site s_i . Le calcul de X_2^i est réalisé de la manière suivante :

$$X_2(s_i) = \sum_{j=1}^{j=p_i} f_{ij}^{sat} \cdot f_{ij}^{int} \cdot \Delta_{ij}^{teinte} \quad \forall s_j \in \eta_i, \quad (4)$$

où Δ_{ij}^{teinte} représente la différence de teinte entre le site s_i et le site voisin s_j . Du fait de la nature circulaire de la teinte, nous calculons la différence de teinte entre deux sites s_i et s_j de la façon suivante :

$$\Delta_{ij}^{teinte} = \begin{cases} \Delta_{ij}^\mu & \text{pour } \Delta_{ij}^\mu \leq 0,5 \\ 1 - \Delta_{ij}^\mu & \text{sinon} \end{cases} \quad (5)$$

où $\Delta_{ij}^\mu = |H(s_i) - H(s_j)|$, $H(s_i)$ étant la valeur de teinte du site s , celle-ci étant comprise entre 0 et 1. Une valeur

de teinte égale à 1 représente un angle de 360° et donc une valeur de 0, 5 représente un angle de 180°. Les sites voisins ayant un contraste important en terme de teinte avec le site s_i vont augmenter le poids de cette seconde contribution à la saillance finale.

Dans la troisième étape, nous étendons le principe de contraste entre couleurs chaudes et froides, au contraste entre couleurs passives et actives. Une couleur est considérée comme étant active si sa valeur de teinte est comprise dans la première moitié du disque de représentation de la teinte, c'est-à-dire, une valeur inférieure à 0, 5 (180°). Ainsi, si la couleur d'un site s_i est active, alors un site s_j avec une couleur passive va contribuer à la saillance du site s_i . Une différence importante en terme de teinte va rendre ce contraste plus saillant. Cette contribution à la saillance du site s_i s'écrit sous la forme :

$$X_3(s_i) = \sum_{j=1}^{j=p_i} f_{ij}^{sat} \cdot f_{ij}^{int} \cdot \Delta_{ij}^{teinte} \quad \forall s_j \in \eta_i, \quad (6)$$

$si \ H(s_i) < 0,5 \text{ et } H(s_j) \geq 0,5$

La quatrième étape constitue la contribution liée au contraste de saturation. Les sites possédant des différences de saturation importantes dans leur voisinage contribuent à la saillance du site s_i de la façon suivante :

$$X_4(s_i) = \sum_{j=1}^{j=p_i} f_{ij}^{sat} \cdot f_{ij}^{int} \cdot \Delta_{ij}^{sat} \quad \forall s_j \in \eta_i, \quad (7)$$

où Δ_{ij}^{sat} est la différence de saturation entre les sites s_i et s_j .

La cinquième étape regroupe les contributions pour le site s_i , à partir des blocs voisins ayant une différence importante en terme d'intensité (contraste d'intensité). Le formule utilisée est similaire à celui de la quatrième étape et s'écrit :

$$X_5(s_i) = \sum_{j=1}^{j=p_i} f_{ij}^{sat} \cdot f_{ij}^{int} \cdot \Delta_{ij}^{int} \quad \forall s_j \in \eta_i, \quad (8)$$

où Δ_{ij}^{int} est la différence d'intensité entre les sites s_i et s_j . Pour chaque site s_i , p_i sites voisins ont contribué à la saillance dans les cinq premières étapes. Les contributions finales sont obtenues en fonction du nombre de voisins pour chaque site s_i :

$$V_\zeta(s_i) = \frac{X_\zeta^i}{p_i} \quad \forall \zeta \in \{1..5\}, \quad (9)$$

où p_i est le nombre de voisins disponibles dans le voisinage 8-connexe.

Les couleurs chaudes constituées de l'intervalle de couleurs rouge, orange et jaune produisent une contribution supplémentaire afin de renforcer leur saillance dans la sixième étape. Cet intervalle de couleur est situé dans les premiers 45° du disque de représentation de la teinte. Cette contribution se formule de la façon suivante :

$$V_6(s_i) = \begin{cases} S(s_i) \cdot I(s)_i & \text{pour } 0 \leq H(s_i) < 0,125 \\ 0 & \text{sinon} \end{cases} \quad (10)$$

la valeur de la teinte variant entre 0 et 1, un angle de 45° correspond à une valeur de 0, 125.

Finalement, la septième étape est constituée de la contribution liée aux sites ayant une couleur fortement saturée et une intensité lumineuse importante. Ces composantes de couleur du site s_i sont combinées afin de déterminer la contribution pour la dernière étape :

$$V_7(s_i) = S(s_i) \cdot I(s_i) \quad (11)$$

La saillance spatiale finale est obtenue en combinant les contributions des sept étapes :

$$S^{SP}(s_i) = \frac{1}{7} \sum_{\zeta=1}^7 V_\zeta(s_i) \quad (12)$$

Cette carte est ensuite normalisée en fonction de la saillance maximale globale obtenue pour chaque image :

$$S^{SP'}(s_i) = S^{SP}(s_i) / S_{max},$$

où S_{max} est la valeur de saillance maximale obtenue pour l'un des blocs de l'image.

3 Saillance temporelle

L'aspect temporel est primordial dans la modélisation de l'attention visuelle. Dans un contexte de recherche visuelle, J. Wolfe [3] a clairement identifié le mouvement comme un attracteur visuel. Une cible en mouvement enfouie dans un ensemble de distracteurs fixe attire l'attention. En outre, une cible fixe enfouie dans un ensemble de distracteurs en mouvement attire l'attention mais dans une moindre mesure. Dans ce contexte d'études, le contraste en mouvement est l'élément déterminant qui attire notre attention visuelle. La cible en contraste de mouvement saute littéralement aux yeux.

De plus, pour la détection de zones saillantes d'une séquence d'images projetées sur un écran, il est intéressant d'avoir à l'esprit les règles en vigueur liées à la capture du film. Les mouvements de caméra influencent clairement la stratégie visuelle de l'observateur. La présence ou non de mouvement permet de hiérarchiser les différents événements. Par ailleurs, la prise de vue est significative du "message" que le metteur en scène souhaite faire passer. Elle incite inconsciemment le téléspectateur à regarder "quelque chose à un endroit particulier".

En conclusion, l'objectif est de déterminer les zones présentant un contraste de mouvement. À partir des données issues de l'estimation du mouvement global et de la segmentation spatio-temporelle [7], il est possible de déterminer le contraste de mouvement pour chaque objet et plus

particulièrement pour chaque tube. Ce contraste de mouvement étant la base de la construction de la saillance temporelle.

3.1 Mouvement dominant

Afin de réaliser la segmentation spatio-temporelle, nous avons premièrement estimé le mouvement global à l'aide d'un modèle affine à six paramètres :

$$\begin{pmatrix} V_x \\ V_y \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (13)$$

L'équation 13 donne le déplacement (V_x, V_y) d'un point à la position (x, y) en fonction de six paramètres liés au mouvement global. Le modèle affine réduit le nombre de mouvements de la caméra à trois types : les translations (t_x et t_y), les rotations (a_2, a_3) et les zooms (a_1, a_4). Nous avons adapté la méthode de Coudray [6, 7] pour estimer ces six paramètres.

Lors de l'estimation du mouvement global, nous avons déterminé les paramètres de translation en localisant le maximum de l'histogramme d'accumulation des vecteurs compensés par les paramètres de déformation. Après étude de tous les pics, une segmentation au sens du mouvement, en plus de l'estimation du mouvement global, est effectuée avec l'hypothèse que chaque pic représente le mouvement d'un objet.

Cette méthode d'estimation du mouvement global possède cependant un léger défaut. En effet, les paramètres de translation du mouvement global sont détectés à l'aide du pic principal dans l'histogramme d'accumulation. Si la séquence vidéo traitée contient un objet uniforme de taille importante, c'est-à-dire, recouvrant plus de la moitié de l'image, les vecteurs de mouvement de cet objet vont alors être identifiés comme le pic principal dans l'histogramme d'accumulation. Afin de résoudre ce problème et d'identifier correctement le mouvement apparent dominant de la séquence, les blocs de chaque objet segmenté situés sur le bord de l'image sont comptabilisés. L'objet possédant le plus grand nombre de blocs situés sur le bord de l'image sera identifié comme le fond de la scène. Le vecteur de translation associé à cet objet sera donc identifié comme le mouvement apparent dominant.

3.2 Mouvement relatif et saillance temporelle

À partir de la connaissance du mouvement apparent dominant \vec{V}_Θ et du déplacement local \vec{V}_{local} pour chaque site (macrobloc du tube situé sur l'image centrale du segment temporel de neuf images), le mouvement relatif \vec{V}_{rel} , exprimé dans le référentiel rétinien est obtenu simplement par la relation suivante :

$$\vec{V}_{rel}(s) = \vec{V}_\Theta(s) - \vec{V}_{local}(s) \quad (14)$$

Le mouvement relatif est nécessaire pour estimer le contraste de mouvement inhérent à un site particulier. Mais

ce n'est pas suffisant de le considérer de cette façon. En effet l'œil est capable de poursuivre des objets en déplacement. Cette faculté liée au mouvement oculaire de poursuite, permet de conserver l'objet suivi dans la fovéa, partie de la rétine présentant la sensibilité spatiale la plus élevée. Par conséquent, considérer directement le mouvement relatif donné par la relation 14 serait réducteur. Il n'est pas correct de dire que plus le mouvement relatif est important, plus la saillance est forte. Il faut prendre en compte la capacité maximale de poursuite de l'œil. S. Daly [8] a montré que la vitesse de poursuite maximale de l'œil pouvait aller jusqu'à 80 deg/sec. Si la vélocité du mouvement relatif est supérieure à la vélocité maximale de poursuite, alors la saillance temporelle est nulle. De plus, celle-ci sera maximale entre $\vec{v}_1 = 20\% \times 30 \text{ deg/sec}$ et $\vec{v}_2 = 30 \text{ deg/sec}$. Pour les vélocités inférieures à \vec{v}_1 et supérieures à \vec{v}_2 , la saillance sera obtenue en fonction d'une droite affine, définie ci-dessous :

$$S^T(s) = \begin{cases} \frac{1}{7} \vec{V}_{rel}(s) & \text{si } 0 \leq \vec{V}_{rel}(s) < \vec{v}_1, \\ 1 & \text{si } \vec{v}_1 \leq \vec{V}_{rel}(s) < \vec{v}_2, \\ \frac{1}{60} \vec{V}_{rel}(s) + \frac{8}{5} & \text{si } \vec{v}_2 \leq \vec{V}_{rel}(s) < \vec{v}_{max} \\ 0 & \text{si } \vec{V}_{rel}(s) \geq \vec{v}_{max} \end{cases} \quad (15)$$

où $\vec{v}_1 = 20\% \times 30 \text{ deg/sec}$, $\vec{v}_2 = 30 \text{ deg/sec}$ et $\vec{v}_{max} = 80 \text{ deg/sec}$. L'indice de saillance temporelle obtenu en fonction de la vélocité temporelle est illustré à la figure 2.

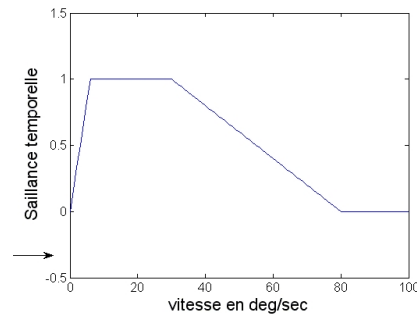


FIG. 2 – Fonction représentant la saillance temporelle en fonction de la vitesse relative.

4 Saillance spatio-temporelle

À partir de la saillance spatiale et de la saillance temporelle, la saillance spatio-temporelle est à déterminer. Les études réalisées par O. Lemeur [9,10] montrent que les observateurs ont tendance à favoriser le centre de l'écran. C'est pourquoi il pondère son modèle de saillance spatiale par une gaussienne bi-dimensionnelle centrée sur l'image. Son étendue spatiale a été optimisée sur une base d'images et sa valeur est de 2,5 degrés visuel. Lors de nos tests,

nous utilisons des séquences vidéos HD (Haute Définition). Celles-ci ont une définition maximale de 1080 lignes par 1920 colonnes. C’est pourquoi, nous avons décidé d’utiliser une gaussienne bi-dimensionnelle centrée sur l’image, dont l’étendue spatiale est égale à 5 degrés visuel. La saillance spatio-temporelle est obtenue en combinant la saillance temporelle et la saillance spatiale pondérées par une gaussienne bi-dimensionnelle de la façon suivante :

$$S^{SP-T}(s) = (S^T(s) + \frac{1}{2}S^{SP'}(s)) \times gauss2D(s), \quad (16)$$

où $gauss2D$ est la gaussienne bi-dimensionnelle d’étendue spatiale égale à 5 degrés visuel. Le mouvement étant l’un des paramètres qui influence le plus l’attention visuelle [11], la pondération de la saillance temporelle est deux fois plus importante que celle de la saillance spatiale.

Finalement, nous obtenons une carte de saillance par groupe de neuf images. Ensuite, on projette cette carte pour les images précédentes et suivantes au sein du segment temporel (neuf images) à l’aide des informations issues de l’estimation du mouvement et de la segmentation spatio-temporelle.

5 Résultats qualitatifs

Les figures 3, 4 et 5 présentent, pour quatre segments temporels successifs des séquences *Tractor*, *New Mobile and Calendar* et *Knightshields*, les différentes cartes de saillance obtenues. Concernant les résultats de la figure 3, on constate que la zone la plus saillante est le tracteur. En effet, le mouvement réel de celui-ci est détecté par notre méthode et ainsi, le tracteur devient la zone la plus saillante. On observe cependant que la saillance du tracteur n’est pas uniforme. Les caractéristiques spatiales du tracteur sont très hétérogènes en terme de couleur et ce sont les roues de couleur rouge (couleur chaude) qui sont les plus saillantes. Concernant les résultats de la figure 4 relatif à la séquence *New Mobile and Calendar*, les zones les plus saillantes sont également les objets en mouvement. Dans les trois premiers segments temporels présentés (43, 44, et 45), le calendrier est une zone saillante alors que pour le segment temporel 46, il ne l’est plus. En effet, on constate que la saillance temporelle basée sur le mouvement ne détecte pas le calendrier comme une zone saillante. Le mouvement de translation du calendrier dans le segment temporel 46 s’estompe et devient quasiment nul, de ce fait, sa saillance temporelle basée sur le mouvement de celui-ci est nulle. Dans les trois premiers segments temporels, la saillance spatio-temporelle du calendrier n’est pas uniforme. En effet, la gaussienne bi-dimensionnelle utilisée pour produire la favorisation du centre de l’écran par les observateurs diminue progressivement la saillance des zones éloignées du centre de l’image. De plus, les chiffres écrits en rouge sur le calendrier sont plus saillants que les zones voisines. Les figurines disposées sur le train en mouvement sont détectées dans les quatre segments tem-

porels. La figurine orange représentant un tigre est détectée comme la zone la plus saillante, du fait de sa position (au centre en bas) et de sa couleur (orange). Pour la dernière séquence testée, les résultats présentés dans la figure 5, semblent corrects également. En effet, la zone la plus saillante est l’homme se déplaçant. Le fond, bien qu’étant immobile (réellement), très riche en informations spatiales est saillant par endroit. Les différents blasons (“*shields*”) sont plus ou moins saillants en fonction de leur couleur. Les blasons possédant des couleurs chaudes (rouge, orange jaune) sont des zones saillantes.



FIG. 3 – Cartes de saillance pour la séquence *Tractor* (segments 13 à 16), avec de haut en bas : images originales, cartes de saillance spatiale, cartes de saillance temporelle et cartes de saillance spatio-temporelle.



FIG. 4 – Cartes de saillance pour la séquence *New Mobile and Calendar* (segments 43 à 46), avec de haut en bas : images originales, cartes de saillance spatiale, cartes de saillance temporelle et cartes de saillance spatio-temporelle.

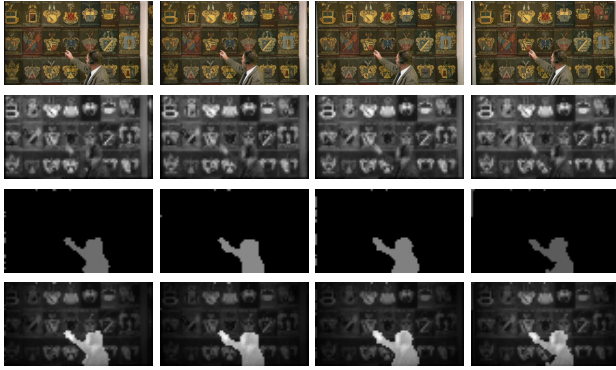


FIG. 5 – Cartes de saillance pour la séquence Knight-shields (segments 28 à 31), avec de haut en bas : images originales, cartes de saillance spatiale, cartes de saillance temporelle et cartes de saillance spatio-temporelle.

6 Conclusion

Cet article a présenté une modélisation de saillance spatio-temporelle. Le calcul de la saillance spatiale est basé sur les couleurs et plus particulièrement sur leur contraste. Comparativement à la détermination de la saillance spatiale, la saillance temporelle est plus facile à calculer car le concept sous-jacent est relativement simple. Une zone temporellement saillante est une zone en contraste de mouvement. Dans l'approche proposée, le contraste de mouvement est déterminé à partir d'une estimation locale (vecteur de mouvement par tube) et d'une estimation globale du mouvement ; la différence, appelée mouvement relatif, indique les zones temporellement saillantes. La saillance finale est obtenue en combinant les saillances spatiale et temporelle. Le centre de l'image en lui-même étant une zone particulière qui attire l'attention visuelle, nous avons utilisé une gaussienne bi-dimensionnelle afin de modéliser ce phénomène lors du calcul de notre saillance spatio-temporelle. Les résultats obtenus en terme de saillance visuelle semblent fiables.

Afin d'évaluer quantitativement les résultats de notre modèle, il serait intéressant de posséder des données réelles en terme de saillance visuelle. Pour cela, on pourra réaliser des expérimentations oculométriques sur les séquences vidéos testées et collecter les données pour construire une référence en terme de saillance visuelle.

Références

- [1] J.M. Wolfe, et T.S. Horowitz, "What attributes guide the deployment of visual attention and how do they do it ?" *Nature Rev. Neuroscience*, vol. 5, pp. 1-7, 2004.
- [2] M. Z. Aziz, et B. Mertsching, "Fast and Robust Generation of Feature Maps for Region-Based Visual Attention." *IEEE Transactions on Image Processing*, vol. 17, no. 5, pp. 633-644, 2008.
- [3] J.M. Wolfe, K.R. Cave, et S.L. Franzel, "Guided search : an alternative to the feature integration model for visual search." *Journal of experimental psychology. Human perception and performance*, vol. 15, no. 3, pp. 419-433, 1989
- [4] J. Itten, "The Elements of Color." New York USA : John Wiley & Sons Inc, 1961.
- [5] F. Mahnke, "Color, Environment, and Human Response." Detroit : Van Nostrand Reinhold, 1996.
- [6] R. Coudray, et B. Besserer, "Global motion estimation for MPEG-encoded streams," in *Proc. IEEE International Conference on Image Processing, ICIP 2004*, Singapore, Republic of Singapore, octobre 2004.
- [7] O. Brouard, F. Delannay, V. Ricordel, et D. Barba, "Robust Motion Segmentation for High Definition Sequences using a Fast Multi-Resolution Motion Estimation based on Spatio-Temporal Tubes" in *Proc. Picture Coding Symposium, PCS 2007*, Lisbonne, Portugal, novembre 2007.
- [8] S. Daly "Engineering Observations from Spatio-velocity and Spatiotemporal Visual Models." In *IS&T/SPIE Conference on Human Vision and Electronic Imaging III.*, SPIE vol. 3299, pp. 180-191, janvier 1998.
- [9] O. Lemeur, P. Le Callet, D. Barba, et D. Thoreau, "A coherent computational Approach to model the bottom-up visual attention " *IEEE transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 28, issue 5, pp. 802-817, mai 2006.
- [10] O. Lemeur, P. Le Callet et D. Barba, "Predicting visual fixations on video based on low-level visual features" *Vision Research*, vol 47, issue 19, pp. 2483-2498, septembre 2007.
- [11] W. Osberger, A.J. Maeder, et N. Bergmann "A Perceptually Based quantization Technique for MPEG Encoding" In *IS&T/SPIE Conference on Human Vision and Electronic Imaging III.*, SPIE vol. 3299, pp. 148-1591, janvier 1998.