



(43) Date de la publication internationale  
22 septembre 2016 (22.09.2016)

WIPO | PCT

(10) Numéro de publication internationale  
**WO 2016/146937 A1**

(51) Classification internationale des brevets :  
G06K 9/62 (2006.01)

(21) Numéro de la demande internationale :  
PCT/FR2016/050574

(22) Date de dépôt international :  
15 mars 2016 (15.03.2016)

(25) Langue de dépôt : français

(26) Langue de publication : français

(30) Données relatives à la priorité :  
15 52155 16 mars 2015 (16.03.2015) FR

(71) Déposants : UNIVERSITE PIERRE ET MARIE CURIE (PARIS 6) [FR/FR]; 4, Place Jussieu, Paris, 75005 (FR). CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE - CNRS - [FR/FR]; 3, rue Michel Ange, Paris, 75016 (FR). INSERM (INSTITUT NATIONAL DE LA SANTÉ ET DE LA RECHERCHE MÉDICALE) [FR/FR]; 101, rue de Tolbiac, Paris, 75013 (FR).

(72) Inventeurs : SIO-HOÏ, Ieng; 28 rue Garibaldi, 93100 Montreuil (FR). RYAD, Benosman; 8 cité des Foyers, 93500 Pantin (FR). BERTRAM, Shi; Hong Kong University of Science and Tech Clear Water Bay, Hong Kong (CN).

(74) Mandataires : CABINET PLASSERAUD et al.; 66, rue de la Chaussée d'Antin, 75440 Paris Cedex 09 (FR).

(81) États désignés (sauf indication contraire, pour tout titre de protection nationale disponible) : AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) États désignés (sauf indication contraire, pour tout titre de protection régionale disponible) : ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), eurasien (AM, AZ, BY, KG, KZ, RU, TJ, TM), européen (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Publiée :

— avec rapport de recherche internationale (Art. 21(3))

(54) Title : METHOD FOR PROCESSING AN ASYNCHRONOUS SIGNAL

(54) Titre : PROCÉDÉ DE TRAITEMENT D'UN SIGNAL ASYNCHRONE

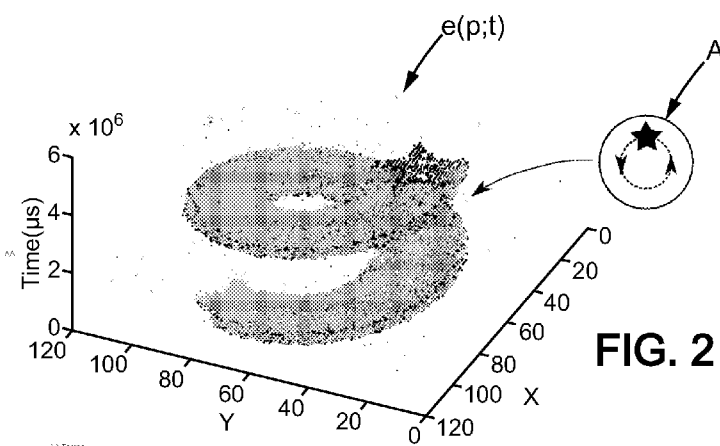


FIG. 2

(57) Abstract : The present invention concerns a method for processing an asynchronous signal produced by a light sensor, the sensor having a pixel matrix disposed opposite a scene, the method comprising: - receiving, from the light sensor, the asynchronous signal (500) comprising, for each pixel of the matrix, successive events from said pixel; - analysing (501) the asynchronous signal using activity profiles as events are received in the asynchronous signal. The activity profile comprises at least, for each pixel (p) of the sensor, an activity value (5(p)) that decreases as a function of the time (t) that has passed since the most recent event (310, 311, 312, 313, 314, 320, 321, 322) among the successive events from said pixel.

(57) Abrégé :

[Suite sur la page suivante]



WO 2016/146937 A1



---

La présente invention concerne un procédé de traitement d'un signal asynchrone produit par un capteur de lumière, le capteur ayant une matrice de pixels disposée en regard d'une scène, le procédé comprenant : - recevoir du capteur de lumière le signal asynchrone (500) comprenant, pour chaque pixel de la matrice, des événements successifs issus dudit pixel; - analyser (501) le signal asynchrone à l'aide de profils d'activité au fur et à mesure de la réception d'événements dans le signal asynchrone. Le profil d'activité comprend au moins, pour chaque pixel (p) du capteur, une valeur d'activité (5(p)) qui décroît en fonction du temps (t) écoulé depuis un événement le plus récent (310, 311, 312, 313, 314, 320, 321, 322) parmi les événements successifs issus dudit pixel.

## **PROCEDE DE TRAITEMENT D'UN SIGNAL ASYNCHRONE**

La présente invention concerne le domaine du traitement de signal asynchrone, notamment pour la détection de formes dans ce signal.

5

Diverses techniques de traitement de signal vidéo ont été historiquement développées. Elles reposent en général sur l'approche tramée traditionnelle.

Il existe des capteurs asynchrones (DVS, ATIS). Ces capteurs peuvent remplacer avantageusement des caméras traditionnelles.

10 Mais le traitement du signal est moins intuitif que dans le monde des caméras tramées. Si certains algorithmes ont été mis au point pour traiter certaines familles de problèmes (ex. flot optique, suivi de formes), d'autres problèmes restent peu ou pas explorés ou sont d'une résolution plus délicate.

Cela est notamment le cas pour la reconnaissance de formes.

15 Afin de détecter des formes et/ou des mouvements, les procédés actuels (i.e. utilisant des données vidéo en provenance de caméras classiques) cherchent à identifier certaines caractéristiques visuelles propres à un ensemble de pixels situés à proximité dans les données vidéo en question.

20 Ces caractéristiques visuelles sont le plus souvent appréhendées par l'homme de l'art comme une information spatiale d'une image (même si cette information spatiale peut être en mouvement).

Dès lors, l'information temporelle de données vidéo est souvent négligée : tout au plus, il peut être recherché une variation/évolution d'une caractéristique visuelle entre deux ou plusieurs images (ou trames) des données vidéo.

25 La négligence de la composante temporelle peut s'expliquer principalement par la technologie d'acquisition habituelle des vidéo : les données vidéo sont le fruit d'une acquisition produisant un grand nombre d'images statiques (ou *frame* en anglais,

trame en français).

Cette technologie d'acquisition des vidéos a conditionné historiquement la manière dont sont traitées ou affichées les vidéos. L'existence de cette succession d'images statiques rend les informations temporelles (ou dynamiques) des données vidéo difficilement manipulables.

S'il est possible d'augmenter le nombre d'images par seconde des données vidéo, il est rare que la fréquence de ces images dépasse 100Hz, tant pour des questions de limites matérielles au niveau de l'acquisition que pour des questions de traitement en temps réel de ces données avec les outils classiques de détection de forme.

10 Il existe ainsi un besoin pour la détection de formes fiable et pouvant être utilisé en exploitant au mieux l'information temporelle des données vidéo.

La présente invention vient ainsi améliorer la situation.

15 Contrairement aux caméras classiques qui enregistrent des images successives à des instants d'échantillonnage réguliers, il a été développé des rétines biologiques inspirées du fonctionnement de l'œil humain. Les rétines biologiques ne transmettent que peu d'informations redondantes sur la scène à visualiser, et ce de manière asynchrone.

Des capteurs de vision asynchrones basés sur événement délivrent des données numériques compressées sous forme d'événements.

20 Une présentation de tels capteurs peut être consultée dans « *Activity-Driven, Event-Based Vision Sensors* », T. Delbrück, et al., Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 2426-2429. Les capteurs de vision basés sur événement ont pour avantage d'enlever de la redondance, de réduire les temps de latence et d'augmenter la gamme dynamique par rapport aux caméras classiques.

25 La sortie d'un tel capteur de vision peut consister, pour chaque adresse de pixel, en une séquence d'événements asynchrones représentatifs des changements de réflectance de la scène au moment où ils se produisent.

Chaque pixel du capteur est indépendant et détecte des changements d'intensité

supérieurs à un seuil depuis l'émission du dernier événement (par exemple un contraste de 15% sur le logarithme de l'intensité). Lorsque le changement d'intensité dépasse le seuil fixé, un événement ON ou OFF est généré par le pixel selon que l'intensité augmente ou diminue (capteurs DVS). Certains capteurs asynchrones  
5 associent les événements détectés à des mesures d'intensité lumineuse (capteurs ATIS).

Le capteur n'étant pas échantillonné sur une horloge comme une caméra classique, il peut rendre compte du séquençement des événements avec une très grande précision temporelle (par exemple, de l'ordre de 1  $\mu$ s). Si on utilise un tel  
10 capteur pour reconstruire une séquence d'images, on peut atteindre une cadence d'images de plusieurs kilohertz, contre quelques dizaines de hertz pour des caméras classiques.

La grande précision temporelle de ces caméras peut permettre d'exploiter au mieux l'information temporelle d'une vidéo.

15 Néanmoins, le traitement des événements en provenance de ces capteurs peut être complexe, car les événements sont des concepts ponctuels en temps (t) et dans l'espace (x,y). Dès lors, leur traitement et leur analyse peuvent être difficiles.

Il existe ainsi un besoin de créer des instruments simples et manipulables pour faire une analyse pertinente d'un signal issu d'un capteur asynchrone.

20 L'analyse doit notamment inclure les dimensions spatiales et la dimension temporelle pour faciliter l'identification de caractéristiques spatio-temporelles dans les événements extrêmement nombreux que délivre un tel capteur, sans perdre la dynamique de celui-ci.

25 À cet effet, la présente invention propose un procédé de traitement d'un signal asynchrone produit par un capteur de lumière, le capteur ayant une matrice de pixels disposée en regard d'une scène, le procédé comprenant :

- recevoir du capteur de lumière le signal asynchrone comprenant, pour chaque pixel de la matrice, des événements successifs issus dudit pixel ;

- analyser le signal asynchrone à l'aide de profils d'activité au fur et à mesure de la réception d'événements dans le signal asynchrone.

Le profil d'activité comprend au moins, pour chaque pixel du capteur, une valeur d'activité qui décroît en fonction du temps écoulé depuis un événement le plus récent  
5 parmi les événements successifs issus dudit pixel.

Le « profil d'activité » d'un pixel peut être vu comme une courbe fonction du temps dont la valeur est représentative, au moins, du temps du dernier événement reçu pour ce pixel (éventuellement filtré sur une polarité donnée).

Les profils d'activité ainsi établis constituent des outils d'analyse qui résument  
10 l'apparition des événements en conservant localement leur structure spatio-temporelle.

Il a été constaté que la morphologie des profils d'activité dénote la présence de certaines formes élémentaires dans la scène observée par le capteur.

Ces profils évoluent au fur et à mesure de la réception des événements  
15 asynchrones et conservent donc la dynamique du capteur.

L'ensemble des profils d'activité constitue ce que l'on peut appeler une « carte de fraîcheur » du capteur.

De nombreuses applications de traitement de signal peuvent être mises au point avec l'analyse utilisant les profils d'activité. De façon non limitative, il est possible de  
20 les utiliser pour :

- la reconnaissance de formes: il est fait référence à une bibliothèque de profils types traduisant des mouvements prédéfinis de formes élémentaires prédéfinies. Si l'analyse révèle un profil d'activité similaire à un profil type de la bibliothèque, une reconnaissance de forme a lieu et une indication de  
25 vitesse de déplacement de cette forme est estimée ;
- estimation de mouvement dans la scène : si un profil d'activité est détecté une première fois puis peu après (ou un profil d'activité très similaire) dans une région proche, une estimation sur la vitesse de déplacement d'un objet dans le champ de vision du capteur peut être déterminée ;

- estimation de profondeur dans une image (stéréo) : si un même profil d'activité (ou similaire) est détecté dans deux images provenant de deux capteurs capturant la même scène, il est possible de considérer que les pixels associés aux profils d'activités correspondent au même point de la scène capturé et ainsi déterminer la profondeur de ce point ;
- etc.

Le « signal asynchrone » peut être l'ensemble des événements provenant d'un capteur rétinien donné ou un sous-ensemble de ces événements (sous-ensemble spatial : limité à certains pixels seulement ; ou/et sous-ensemble temporel : limité à une période de temps donné).

En outre, le profil d'activité peut décroître exponentiellement en fonction du temps écoulé depuis un événement le plus récent parmi les événements successifs issus dudit pixel.

Cette décroissance exponentielle peut permettre de mieux mettre en évidence les derniers événements reçus par le capteur.

Dans un mode de réalisation particulier, le profil d'activité peut être en outre en fonction du temps écoulé depuis un événement antérieur à l'événement le plus récent parmi les événements successifs issus dudit pixel.

Il est même possible de prendre en compte tous les événements reçus pour ce pixel afin de déterminer le profil d'activité associé.

La présente invention propose également un procédé de reconnaissance de formes spécialement adapté aux capteurs rétiniens, et utilisant le traitement du signal asynchrone précédemment présenté.

La présente invention peut également viser un procédé de reconnaissance de formes comprenant :

/a/ l'exécution du procédé de traitement d'un signal asynchrone présenté

précédemment ;

/b/ au fur et à mesure de la réception des événements, pour chaque pixel courant dudit capteur :

5                   /b1/ détermination d'un contexte pour ledit pixel courant, ledit contexte étant défini comme un ensemble de profils d'activité de pixels situés à moins d'une distance prédéterminée dudit pixel courant,

10                   /b2/ détermination d'un contexte type parmi un ensemble de contextes types prédéterminés, le contexte type déterminé ayant, parmi l'ensemble des contextes types, une distance au contexte déterminé à l'étape /b1/ minimum,

/c/ détermination d'une signature fonction d'un nombre d'occurrences de détermination des contextes types à l'étape /b2/;

/d/ détermination d'une forme par une comparaison de ladite signature avec une base de signature types.

15       Les distances peuvent être des distances au sens mathématique. Ainsi, les distances peuvent être des distances euclidiennes, des distances de *Manhattan*, des distances de *Minkowski*, des distances de *Tchebychev* ou toutes autres distances.

20       Le plus souvent, il est possible de représenter un « contexte » comme une surface dans un espace à trois dimensions (deux axes représentatifs des coordonnées des pixels et un axe sans dimension (fonction d'un temps)).

25       Ce contexte permet d'appréhender l'environnement proche de l'événement considéré (i.e. courant) spatialement et temporellement. Un contexte peut être vu comme un ensemble de valeurs associées à des coordonnées des pixels situés à moins d'une distance prédéterminée d'un pixel dont provient l'événement courant (dit « pixel courant »), relativement au pixel courant.



Le plus souvent, il est possible de représenter un « ensemble de composantes » graphiquement comme un « histogramme ». La terminologie « histogramme » ou « signature » est également utilisée dans la suite de la description.

- 5 Dans un autre mode de réalisation, il est possible de prendre en compte un modèle hiérarchique pour les contextes types utilisés.

La présente invention peut ainsi viser un procédé de reconnaissance de formes comprenant (un modèle hiérarchique de contextes types étant définis, chaque contexte type étant associé à une pluralité de niveau du modèle hiérarchique) :

- 10 /a/ utilisation des événements du signal asynchrone du capteur comme événements courants et utilisation des contextes types du premier niveau du modèle hiérarchique comme contextes types courants ;
- /b/ exécution du procédé de traitement d'un signal asynchrone comme indiqué précédemment avec les événements courants ;
- 15 /c/ au fur et à mesure de la réception des événements courants, pour chaque pixel courant dudit capteur :
- /c1/ détermination d'un contexte pour ledit pixel courant, ledit contexte étant défini comme un ensemble de profils d'activité utilisés à l'étape
- /b/ pour des pixels situés à moins d'une distance prédéterminée dudit
- 20 pixel courant,
- /c2/ détermination d'un contexte type ( $C_k$ ) parmi les contextes types courants (504), le contexte type déterminé ayant, parmi l'ensemble des contextes types courants, une distance au contexte déterminé à l'étape /c1/ minimum,
- 25 /c3/ génération d'un événement fonction du contexte type identifié à l'étape /c2/ ;
- /d/ si un niveau du modèle hiérarchique n'a pas été utilisé :
- utilisation des événements générés à l'étape /c3/ comme événements courants,

- utilisation des contextes types d'un niveau du modèle hiérarchique non encore utilisé comme contextes types courants ;
- réexécution des étapes /b/ à /e/ ;

/e/ si tous les niveaux du modèle hiérarchique ont été utilisés :

- 5
  - détermination (521) d'une signature fonction d'un nombre d'occurrences de détermination des contextes types à la dernière occurrence de l'étape /c2/ ;
  - détermination (524) d'une forme par une comparaison de ladite signature avec une base de signatures types.
- 10 Ces niveaux hiérarchiques permettent de procéder à une reconnaissance de proche en proche : le premier niveau hiérarchique permet d'identifier des formes très basiques, les niveaux hiérarchiques d'ordre supérieur permettant d'identifier des formes plus complexes à partir des niveaux inférieurs.
- 15 Dans un mode de réalisation particulier, la détermination d'un contexte prend en compte, séparément, les événements ayant des polarités différentes.

Par exemple, la polarité peut correspondre au fait qu'une intensité de pixel varie (ex. +1 en cas d'augmentation ou -1 en cas de diminution). La polarité peut également correspondre au contexte type identifié pour le niveau hiérarchique
- 20 immédiatement inférieur.

Le fait de prendre en compte séparément les événements ayant des polarités différentes permet d'augmenter la pertinence des contextes déterminés.

Dans ce cas, un contexte peut être vu comme un ensemble de valeurs associées :

  - à des coordonnées des pixels situés à moins d'une distance prédéterminée
- 25 d'un « pixel courant », relativement au pixel courant ; et
  - à une polarité donnée.

Avantageusement, la distance utilisée à l'étape /e/ est une distance de *Bhattacharyya* ou une distance normalisée.

- 5 Un programme informatique, mettant en œuvre tout ou partie du procédé décrit ci-avant, installé sur un équipement préexistant, est en lui-même avantageux.

Ainsi, la présente invention vise également un programme informatique comportant des instructions pour la mise en œuvre du procédé précédemment décrit, lorsque ce programme est exécuté par un processeur.

- 10 Ce programme peut utiliser n'importe quel langage de programmation (par exemple, un langage-objet ou autre), et être sous la forme d'un code source interprétable, d'un code partiellement compilé ou d'un code totalement compilé.

La figure 5 décrite en détail ci-après peut former l'organigramme de l'algorithme général d'un tel programme informatique.

15

D'autres caractéristiques et avantages de l'invention apparaîtront encore à la lecture de la description qui va suivre. Celle-ci est purement illustrative et doit être lue en regard des dessins annexés sur lesquels :

20

- la figure 1 est un schéma synoptique d'un capteur asynchrone de lumière de type ATIS ;

- la figure 2 est un diagramme montrant des événements générés par un capteur asynchrone placé en regard d'une scène comportant une étoile tournante ;

25

- les figures 3a et 3b sont des exemples de représentation d'un « signal d'activité » pour les événements d'un pixel donné ;

- la figure 3c est un exemple d'une surface 3D représentant une carte de « fraîcheur » donnée selon un mode de réalisation de l'invention ;

- la figure 3d est un exemple d'une surface 3D représentant un contexte selon

un mode de réalisation de l'invention ;

- la figure 4 illustre la survenance d'événements dans le flux d'événements lorsque des cartes à jouer sont retournées ;

5       - la figure 5 est un exemple d'ordinogramme d'un mode de réalisation possible de l'invention ;

- les figures 6a à 6c sont des exemples de contextes types dans un mode de réalisation possible de l'invention ;

- la figure 7 illustre un dispositif pour la mise en œuvre d'un mode de réalisation selon l'invention ;

10       - les figures 8a à 8c sont des exemples de performances de détection de formes pour différentes formes à reconnaître.

La figure 1 illustre le principe de l'ATIS.

15       Un pixel 101 de la matrice constituant le capteur comporte deux éléments photosensibles 102a, 102b, tels que des photodiodes, respectivement associés à des circuits électroniques de détection 103a, 103b.

Le capteur 102a et son circuit 103a produisent une impulsion  $P_0$  lorsque l'intensité lumineuse reçue par la photodiode 102a varie d'une quantité prédéfinie.

20       L'impulsion  $P_0$  marquant ce changement d'intensité déclenche le circuit électronique 103b associé à l'autre photodiode 102b. Ce circuit 103b génère alors une première impulsion  $P_1$  puis une deuxième impulsion  $P_2$  dès qu'une quantité de lumière donnée (nombre de photons) est reçue par la photodiode 102b.

L'écart temporel  $\delta t$  entre les impulsions  $P_1$  et  $P_2$  est inversement proportionnel à l'intensité lumineuse reçue par le pixel 101 juste après l'apparition de l'impulsion  $P_0$ .

25       L'information asynchrone issue de l'ATIS comprend deux trains impulsions combinées pour chaque pixel (104) : le premier train des impulsions  $P_0$  indique les instants où l'intensité lumineuse a changé au-delà du seuil de détection, tandis que

le second train se compose des impulsions  $P_1$  et  $P_2$  dont l'écart temporel  $\delta t$  indique les intensités lumineuses, ou niveaux de gris, correspondants.

Un événement  $e(\mathbf{p}, t)$  issu d'un pixel 101 de position  $\mathbf{p}$  dans la matrice de l'ATIS comporte alors deux types d'information: une information temporelle donnée par la position de l'impulsion  $P_0$ , donnant l'instant  $t$  de l'événement, et une information de niveau de gris donnée par l'écart temporel  $\delta t$  entre les impulsions  $P_1$  et  $P_2$ .

On peut alors placer les événements provenant des pixels dans une représentation espace-temps à trois dimensions telle que celle présentée sur la figure 2. Sur cette figure, chaque point  $\mathbf{p}$  repère un événement  $e(\mathbf{p}, t)$  engendré de manière asynchrone à un instant  $t$  au niveau d'un pixel  $\mathbf{p}$  du capteur, de position  $\begin{pmatrix} x \\ y \end{pmatrix}$ , par le mouvement d'une étoile tournant à vitesse angulaire constante comme schématisé dans l'encadré A. La majeure partie de ces points se répartissent à proximité d'une surface de forme générale hélicoïdale. En outre, la figure montre un certain nombre d'événements à distance de la surface hélicoïdale qui sont mesurés sans correspondre au mouvement effectif de l'étoile. Ces événements sont du bruit d'acquisition.

Les événements  $e(\mathbf{p}, t)$  peuvent alors être définis par l'ensemble des informations suivantes :

$$e(\mathbf{p}, t) = \begin{cases} \mathbf{p} \in C \subset \mathbb{R}^2 \\ pol \\ I(\mathbf{p}, t) \end{cases}$$

avec  $C$  le domaine spatial du capteur,  $pol$  la polarité représentant la direction du changement de luminance (ex. 1 pour une augmentation ou -1 pour une diminution) et  $I(\mathbf{p}, t)$  le signal d'intensité lumineuse du point  $\mathbf{p}$ .

Le signal d'intensité lumineuse peut être ainsi l'ensemble des trains d'impulsions combinées 104 tel que décrit en Figure 1. Si  $t_u$  représente le temps de survenance de l'événement, et l'écart entre  $t_{e+,u}$  et  $t_{e-,u}$  représente une valeur inversement proportionnelle à la variation d'intensité, il est possible de coder l'intensité à l'aide de trois Dirac  $\delta$  de telle sorte que  $I(\mathbf{p}, t) = \delta(t - t_u) + \delta(t - t_{e+,u}) + \delta(t - t_{e-,u})$ . Le signal d'intensité du pixel situé à la coordonnée  $\mathbf{p}$  permet alors de coder

temporellement les informations de luminance.

Il est possible de noter le  $i^{\text{ème}}$  événement d'un capteur comme  $ev(i)$ ,  $ev(i)$  étant alors défini par l'ensemble des informations suivantes :

$$ev(i) = \begin{cases} \mathbf{p}_i \in C \subset \mathbb{R}^2 \\ pol_i \\ t_i \\ I_i(\mathbf{p}_i, t_i) \end{cases}$$

avec  $C$  le domaine spatial du capteur,  $\mathbf{p}_i$  le point du capteur concerné par le  $i^{\text{ème}}$  événement,  $pol_i$  la polarité représentant un type d'événements (par exemple, la direction du changement de luminance pour le  $i^{\text{ème}}$  événement, ex. 1 pour une augmentation ou -1 pour une diminution),  $t_i$  le temps de survenance du  $i^{\text{ème}}$  événement et  $I_i(\mathbf{p}, t)$  le signal d'intensité lumineuse du point  $\mathbf{p}_i$  (si cette valeur est disponible).

- 10 Afin de manipuler de manière plus simple les événements, il est possible de définir pour chaque pixel  $\mathbf{p}$  et pour chaque polarité  $pol$  une fonction  $S$  représentative d'une « fraîcheur » des événements,  $S(\mathbf{p}, pol)$  étant fonction au moins du temps  $t$  de survenance du dernier événement pour ce pixel et ayant cette polarité  $pol$ .

Il est possible de définir la fonction  $S$  comme la somme, pour chaque événement

$$15 \quad ev_p(i) = \begin{cases} \mathbf{p}_i = \mathbf{p} \\ pol_i = pol \\ t_p \end{cases} \text{ d'un pixel } \mathbf{p} \text{ donné et pour une polarité } pol \text{ donnée, à un instant}$$

$t$  donné, de la fonction primitive  $S_{prim}(\mathbf{p}, pol, t) = \begin{cases} \max(h + \theta \cdot (t_p - t), 0) & \text{si } t \geq t_p, \\ 0 & \text{sinon} \end{cases}$

étant une valeur prédéterminée et  $\theta$  étant un facteur prédéterminé correspondant à la vitesse de la décroissance de, la fonction primitive.

- 20 La « somme » de la fonction primitive peut être également vue mathématiquement comme une convolution :

- de la fonction primitive  $S_{prim}(\mathbf{p}, t) = \begin{cases} \max(h - \theta \cdot t, 0) & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases}$  (ou plus généralement de toute fonction décroissante),

- avec un signal comportant un Dirac pour chaque temps  $t_p$  auquel un

$$\text{événement } ev_p(i) = \begin{cases} p_i = p \\ pol_i = pol \\ t_p \end{cases} \text{ est survenu.}$$

A titre d'illustration, la figure 3a montre trois signaux d'activité possibles  $t \rightarrow S$ , pour trois pixels  $p_1$ ,  $p_2$  et  $p_3$  du capteur (et pour une valeur de polarité  $pol$  donnée).

- En l'absence d'événements, la valeur de  $S(p_1, t)$ ,  $S(p_2, t)$  ou  $S(p_3, t)$  est nulle.
- 5 Néanmoins, lors de la survenance d'un événement de polarité  $pol$  (par exemple, 310) au niveau du pixel  $p_1$ ,  $S(p_1, t)$  prend une valeur seuil prédéterminé (ici  $h$ , cette valeur  $h$  pouvant être unitaire).

La valeur du signal d'activité  $S(p_1, t)$  décroît alors progressivement après cet événement pour tendre vers 0.

- 10 Il en va de même pour l'événement 311 pour le pixel  $p_1$ , pour l'événement 312 pour le pixel  $p_2$ , ou pour l'événement 313 / 314 pour le pixel  $p_3$ ,

Si la décroissance du signal d'activité  $S$  est ici linéaire, il est possible de prévoir tout type de décroissance comme une décroissance exponentielle :

$$h \cdot e^{-\frac{t-t_p}{\tau}}$$

- 15 Cette décroissance exponentielle peut être illustrée par la figure 3b (voir la courbe 3b et l'événement 320).

Par ailleurs, il est possible que, lors de la survenance d'un événement pour le pixel considéré (ex.  $p_4$  ici), la valeur de la fonction  $S$  ne soit pas négligeable par rapport à la valeur de  $h$  (ex. l'événement 321 est temporellement proche de l'événement 322).

- 20 Dans un mode de réalisation, lors de la survenance de l'événement ultérieur 322, la valeur du signal d'activité  $S$  peut être fixée à la somme (éventuellement pondérée) de la valeur courante de  $S$  juste avant l'événement 322 (i.e.  $h_0$ ) et de  $h$ . Ainsi, la décroissance de la courbe  $S$  partira de la valeur  $h+h_0$  comme le montre la figure 3b.
- Par ailleurs, il est possible de prévoir que la valeur de  $h+h_0$  soit capée à une valeur
- 25  $h_1$  prédéterminée (i.e.  $\min(h_1, h+h_0)$ )

Dans un autre mode de réalisation, lors de la survenance de l'événement ultérieur

322, la valeur de la courbe  $S$  est fixée à la valeur  $h$  quelle que soit la valeur de  $h_0$  (i.e. les événements antérieurs au dernier événement (i.e. l'événement ultérieur) sont ignorés). Dans cet autre mode de réalisation, il est possible de définir un temps dit « temps de dernier événement » défini comme suit :

$$T(\mathbf{p}, pol, i) = \max(t_j) \mid j < i$$

5 ou

$$T(\mathbf{p}, pol, t) = \max(t_j) \mid t_j < t$$

avec  $t_j$  les temps d'événements survenant pour le pixel pour un pixel  $\mathbf{p}$  avec la polarité  $pol$ .

Conceptuellement,  $\mathbf{p} \rightarrow T(\mathbf{p}, pol, t)$  définit une carte des temps des derniers événements de même polarité survenus temporellement juste avant un temps de  
10 référence (i.e.  $t$ ).

On peut alors définir, dans cet autre mode de réalisation,  $\mathbf{p} \rightarrow S(\mathbf{p}, pol, t)$  comme étant une fonction de cet ensemble de temps  $T(\mathbf{p}, pol, t)$ .

Par exemple,  $\mathbf{p} \rightarrow S(\mathbf{p}, pol, t)$  :

$$\mathbf{p} \rightarrow S(\mathbf{p}, pol, t) : \begin{cases} h \cdot e^{-\frac{t-T(\mathbf{p}, pol, t)}{\tau}} & \text{si } t \geq T(\mathbf{p}, pol, t) \\ 0 & \text{sinon} \end{cases}$$

avec  $\tau$  et  $h$  une constante temporelle prédéterminée ( $S$  peut être toute fonction  
15 décroissante avec le temps  $t$  sur un intervalle comprenant comme borne inférieure  $T(\mathbf{p}, pol, t)$ ).

La création d'une carte  $S$  de pixels représentative de la « fraîcheur » d'événements de ces pixels est avantageuse, car elle permet une représentation continue et simple de concepts discontinus (i.e. les événements). Cette carte créée permet de  
20 transformer la représentation des événements dans un domaine simple d'appréhension.

Dès lors, sa création simplifie la manipulation et la comparaison des événements.



Une fois la carte de pixels  $\mathbf{p} \rightarrow S(\mathbf{p}, pol, t)$  déterminée, il est possible de réaliser un graphique en 3D de l'amplitude de  $S$  en fonction des coordonnées  $\mathbf{p}$ , pour un temps  $t$  et pour une valeur de polarité  $pol$  fixés (voir Figure 3c).

Bien entendu,  $S$  peut comprendre  $N$  jeux de valeurs séparés (i.e.  $(\mathbf{p}, pol) \rightarrow S(\mathbf{p}, pol, t)$ ), une pour chaque valeur de polarité  $pol$  possible (s'il existe  $N$  valeurs de polarité possibles).

On appelle  $S_p$  le « contexte » du pixel  $\mathbf{p}$  l'ensemble des valeurs de  $(\mathbf{q}, pol) \rightarrow S(\mathbf{q}, pol, t)$  pour les pixels  $\mathbf{q}$  à proximité du pixel  $\mathbf{p}$  (i.e. situé à une distance prédéterminée du pixel  $\mathbf{p}$ , ex. distance au sens mathématique, par exemple dans un carré de côté  $2R+1$  centré sur le pixel  $\mathbf{p}$ ). Afin de représenter visuellement le contexte (comprenant plusieurs valeurs de polarité possibles), il est possible de juxtaposer plusieurs représentations de  $\mathbf{q} \rightarrow S(\mathbf{q}, pol, t)$  pour les différentes valeurs de  $pol$  possibles.

Par exemple, la figure 3d peut être représentative du « contexte » du pixel  $\mathbf{p}$  : le plateau 351 est alors représentatif d'une valeur de la fonction  $\mathbf{q} \rightarrow S(\mathbf{q}, 1, t)$  et le plateau 352 est représentatif d'une valeur de la fonction  $\mathbf{q} \rightarrow S(\mathbf{q}, -1, t)$  (bien entendu, les plateaux peuvent avoir toutes formes, le choix du plateau permettant simplement de faciliter la visualisation). De même, le point 353 et 354 correspond à une valeur de la fonction  $S$  pour un même pixel mais pour une polarité différente.

Il est possible de définir un contexte  $S_p$  pour tout pixel  $\mathbf{p}$ .

Afin de pouvoir caractériser les différents contextes possibles, il est possible de définir des contextes dits « types ».

Ces contextes types peuvent être prédéterminés ou peuvent être déterminés sur la base de l'algorithme fourni par le document *D. Ballard and J. Jehee, « Dynamic coding of signed quantities in cortical feedback circuits » Frontiers in Psychology, vol. 3 n°254, 2012* ou en utilisant une autre méthode (méthode des « k-means » par exemple).

Pour chaque contexte  $S_p$  identifié pour un pixel  $p$ , il est possible d'associer au pixel  $p$  un contexte type correspondant au contexte type le plus proche de  $S_p$ . La distance entre le contexte  $S_p$  et un contexte type peut être, par exemple, déterminé en calculant une somme des distances euclidiennes entre les valeurs du contexte  $S_p$  et du contexte type pour des mêmes pixels  $p$  et des mêmes polarités  $pol$ . La distance peut être également fonction de la somme des distances euclidiennes au carré.

Si la distance calculée est au-dessus d'un certain seuil prédéterminé, il est possible de n'associer aucun contexte type au pixel  $p$ .

Ces contextes types  $\{C_k\}$  sont définis sur un domaine spatial limité comme précédemment évoqué pour les contextes  $S_p$  (ex. sur des carrés de côté  $2R + 1$ ).

Il est également possible de définir plusieurs niveaux hiérarchiques de contextes types, chaque niveau hiérarchique  $m$  définissant une pluralité de contextes types  $\{C_{k,m}\}$ . L'intérêt de tels niveaux hiérarchiques est détaillé au regard de la description de la figure 5.

A titre d'illustration, la figure 6a illustre un premier niveau hiérarchique de contextes types possibles (i.e. taille 5x5 pixels pour une polarité de 1 (zone de gauche de chacun des contextes types) ou pour une polarité de -1 (zone de droite de chacun des contextes types)).

De plus, à un instant  $t$  donné, et après identification des contextes types  $\{C_k\}$  (ou plus généralement  $\{C_{k,m}\}$  pour un niveau hiérarchique  $m$  fixé) associés à chaque pixel  $p$ , il est possible de calculer un nombre d'occurrences de chacun des contextes types pour l'ensemble des pixels  $p$  possibles.

Ces nombres d'occurrences calculés permettent de créer des signatures / histogrammes  $\mathcal{H}$  caractérisant le flux d'événements (en ordonné, le nombre d'occurrences, en abscisse, l'indice du contexte type identifié).

La caractérisation du flux d'événements permettant la reconnaissance de forme peut aussi utiliser d'autres méthodes telles des Echo-State Networks ou des Réseaux de neurones récurrents.

À titre d'illustration, la figure 4 illustre la survenance d'événements « ON » (points blancs) et d'événements « OFF » (points noirs) lorsque des cartes à jouer comportant un trèfle, un carreau, un cœur, et un pique sont retournées. Les histogrammes respectivement associés à chacun de flux d'événements (a), (b), (c) et (d) sont également affichés en dessous de ces flux d'événements : ces histogrammes constituent une sorte de « signature » des formes à reconnaître dans le flux d'événements.

Il est possible de construire ces histogrammes durant un nombre prédéterminé d'entraînements (i.e. génération d'histogrammes par la présentation d'un flux d'événements représentatifs d'une même forme) : ainsi, il est possible de déterminer un « histogramme type » en moyennant les histogrammes obtenus durant l'entraînement pour une même forme et/ou un même mouvement.

Une fois ces histogrammes types déterminés, il est alors possible de déterminer à nouveau un histogramme courant à partir d'un flux d'événements et de comparer cet histogramme avec les histogrammes types déterminés durant la phase d'entraînement.

L'histogramme type ayant une distance la plus proche avec l'histogramme courant peut alors permettre d'identifier la forme correspondant au flux d'événements.

Il est en outre possible que les k-plus proches histogrammes types soient retournées (éventuellement avec une note correspondant à leur proximité).

La distance entre deux histogrammes  $\mathcal{H}_1$  et  $\mathcal{H}_2$  peut être calculée comme une distance mathématique entre deux vecteurs ayant pour coordonnées les nombres d'occurrences pour chacun des contextes types :

$$d(\mathcal{H}_1; \mathcal{H}_2) = \|\mathcal{H}_1 - \mathcal{H}_2\|$$

Il est également possible de calculer une distance normalisée comme suit :

$$d^N(\mathcal{H}_1; \mathcal{H}_2) = \left\| \frac{\mathcal{H}_1}{\text{card}(\mathcal{H}_1)} - \frac{\mathcal{H}_2}{\text{card}(\mathcal{H}_2)} \right\|$$

avec  $\text{card}(\mathcal{H}_j)$  le nombre de contextes types (i.e. barre verticale) de l'histogramme  $\mathcal{H}_j$ .

La distance de *Bhattacharyya* peut également être utilisée en remplacement de la distance classique :

$$d^B(\mathcal{H}_1; \mathcal{H}_2) = -\ln \sum_i \left( \left\| \frac{\mathcal{H}_1(i)}{\text{card}(\mathcal{H}_1)} \cdot \frac{\mathcal{H}_2(i)}{\text{card}(\mathcal{H}_2)} \right\| \right)$$

avec  $\mathcal{H}_j(i)$  le nombre d'occurrences du  $i^{\text{ème}}$  contexte type de l'histogramme  $\mathcal{H}_j$ .

Toute autre distance mathématique peut également être utilisée.

- 5 Dès lors, il est possible de considérer que la forme correspondant à l'histogramme type associé est survenue dans le flux d'événements.

Grâce à cet histogramme type, une reconnaissance de forme est ainsi possible.

- 10 La figure 5 est un exemple d'ordinogramme d'un mode de réalisation possible de l'invention.

- Sur réception d'un événement  $i$  d'un flux d'événement 500 provenant d'un capteur asynchrone et associé à temps  $t_i$ , il est possible de déterminer ou de mettre à jour (étape 501) les valeurs de  $S(\mathbf{p}, \text{pol}, t_i)$  pour chaque pixel  $\mathbf{p}$  du capteur et pour chaque valeur de  $\text{pol}$ , comme cela est indiqué précédemment : cette étape permet de créer  
15 ou de mettre à jour la carte de « fraîcheur » du capteur.

Pour l'étape 501, la constante de temps utilisée pour la décroissance de la fonction primitive de  $S$  est notée  $\tau_1$ . Ainsi, nous pouvons, par exemple, avoir :

$$S(\mathbf{p}, \text{pol}, t_i) = h. e^{-\frac{t_i - T(\mathbf{p}, \text{pol}, t_i)}{\tau_1}}$$

- Une fois cette détermination effectuée, il est possible, pour chaque pixel  $p$  du capteur, d'extraire un contexte  $S_p$  (étape 502) de la carte de fraîcheur précédemment  
20 calculée à l'étape 501 : cette extraction permet d'isoler certaines valeurs de  $S(\mathbf{q}, \text{pol}, t_i)$  pour les pixels  $\mathbf{q}$  à proximité spatiale  $N_p$  du pixel  $p$  considéré et pour une valeur de polarité  $\text{pol}$  donnée. À titre d'illustration,  $N_{p-1}$  peut définir un carré ou un rectangle centré autour de la position spatiale  $\mathbf{p}$  considérée. La proximité spatiale  $N_{p-1}$  peut être définie afin que les contextes ainsi extraits soient de dimensions

égales aux dimensions des contextes types du premier niveau hiérarchique 503 ( $C_{k-1}$ ), afin d'être comparées à ces derniers.

5 A titre d'illustration, la figure 6a illustre un premier niveau hiérarchique de contextes types possibles (i.e. taille 5x5 pixels pour une polarité de 1 (zone de gauche de chacun des contextes types) ou pour une polarité de -1 (zone de droite de chacun des contextes types)).

La comparaison du contexte déterminé  $S_{p-1}$  et associé au pixel  $p$  avec les contextes types possibles  $\{C_{k-1}\}$  du premier niveau hiérarchique permet d'identifier le contexte type le plus proche de  $S_{p-1}$  (étape 504) comme indiqué précédemment.

10 Cette identification du contexte type le plus proche  $C_{k-1prox}$  parmi avec les contextes types possibles  $\{C_{k-1}\}$  du premier niveau hiérarchique permet de générer, pour le pixel  $p$ , un événement  $ev_1$  indiquant le contexte type associé à ce pixel  $p$  (étape 505) :

$$ev_1(p) = \begin{cases} t \\ C_{k-1prox} \end{cases}$$

avec  $t$  le temps courant.

15 Si aucun contexte type le plus proche n'est identifié (voir supra), aucun événement n'est généré.

La génération de ces événements  $ev_1(p)$  constitue également un flux asynchrone d'événements (506).

20 Dès lors, il est possible de traiter ces événements 506 de manière similaire aux événements 500 en provenance du capteur.

Ainsi, sur réception de chaque événement  $j$  du flux d'événement 506 et associé à temps  $t_j$ , il est possible de déterminer ou de mettre à jour (étape 511) les valeurs d'une nouvelle carte de « fraîcheur » (i.e. carte pour le deuxième niveau hiérarchique) ayant comme valeur  $(p, pol) \rightarrow S_2(p, pol, t_j)$  pour chaque pixel  $p$  du capteur et pour chaque valeur de  $pol$  (sachant que pour cette étape 511 les valeurs de  $pol$  sont représentatives du contexte type identifié lors de l'étape 504 pour le

25

premier niveau hiérarchique, i.e. 4 contextes types possible dans le cadre de la figure 6a, chaque contexte type ayant deux polarités), comme cela est indiqué précédemment.

Pour l'étape 511, la constante de temps utilisée pour la décroissance de la fonction primitive est notée  $\tau_2$  avec  $\tau_2 \geq \tau_1$ . Ainsi, nous pouvons, par exemple, avoir :

$$S_2(\mathbf{p}, t_j) = h. e^{-\frac{t_j - T(\mathbf{p}, t_j)}{\tau_2}}$$

Une fois cette détermination effectuée, il est possible, pour chaque pixel  $p$  du capteur, d'extraire un contexte  $S_{p,2}$  (étape 512) de la carte de fraîcheur précédemment calculée à l'étape 511 : cette extraction permet d'isoler certaines valeurs de  $S_2(\mathbf{q}, pol, t_i)$  pour les pixels  $\mathbf{q}$  à proximité spatiale  $N_{p,2}$  du pixel  $p$  considéré et pour une valeur de polarité  $pol$  donnée. À titre d'illustration,  $N_{p,2}$  peut définir un carré ou un rectangle centré autour de la position spatiale  $\mathbf{p}$  considérée. La proximité spatiale  $N_{p,2}$  peut être définie afin que les contextes ainsi extraits soient de dimensions égales aux dimensions des contextes types du deuxième niveau hiérarchique 513 ( $C_{k,2}$ ), afin d'être comparées à ces derniers.

A titre d'illustration, la figure 6b illustre un deuxième niveau hiérarchique de huit contextes types possibles.

Si chaque contexte type du premier niveau hiérarchique possède deux zones (une zone pour les événements de polarité -1 (OFF) et une zone pour les événements de polarité 1 (ON), voir Figure 6a), les contextes types du deuxième niveau hiérarchique possèdent autant de zones (i.e. 4) que de contexte types dans le niveau hiérarchique immédiatement inférieur (i.e. le premier niveau hiérarchique, ici, voir Figure 6b).

Par ailleurs, les contextes types du deuxième niveau hiérarchique peuvent être tels que la proximité spatiale  $N_{p,2}$  définisse une zone plus grande que celle définie par la proximité spatiale  $N_{p,1}$  (ex. si  $N_{p,1}$  est définie par un carré de côté  $2R_1+1$  alors si  $N_{p,2}$  peut être définie par un carré de côté  $2R_2+1$  avec  $R_2 \geq R_1$ ).

La comparaison du contexte déterminé  $S_{p,2}$  et associé au pixel  $\mathbf{p}$  avec les contextes types possibles  $\{C_{k,2}\}$  du premier niveau hiérarchique permet d'identifier le

contexte type le plus proche de  $S_{p,2}$  (étape 514) comme indiqué précédemment.

Cette identification du contexte type le plus proche  $C_{k\_2prox}$  parmi avec les contextes types possibles  $\{C_{k\_2}\}$  du deuxième niveau hiérarchique permet de générer, pour le pixel  $\mathbf{p}$ , un événement  $ev_2$  indiquant le contexte type associé à ce pixel  $\mathbf{p}$  (étape 515) :

$$ev_2(\mathbf{p}) = \begin{cases} t \\ C_{k\_2prox} \end{cases}$$

avec  $t$  le temps courant.

Si aucun contexte type le plus proche n'est identifié (voir supra), aucun événement n'est généré.

La génération de ces événements  $ev_2(\mathbf{p})$  constitue également un flux asynchrone d'événements (516).

Dès lors, il est possible de traiter ces événements 516 de manière similaire aux événements 506 : s'il existe un niveau hiérarchique supérieur (test 517, sortie OK, par exemple le niveau hiérarchique de niveau trois représenté par la figure 6c (6(c)(1) et 6(c)(2))), il est possible de réitérer les étapes 511 à 515 en prenant en entrée les événements générés par le niveau hiérarchique immédiatement inférieur et en utilisant les contextes types de ce niveau hiérarchique supérieur (voir 513).

S'il n'existe pas de niveau hiérarchique supérieur, il est possible de compter le nombre d'occurrences des contextes types (513) identifiés pour tous les pixels  $\mathbf{p}$  et pour un temps  $t$  fixé. Comme expliqué ci-dessus, ce comptage permet la détermination d'un histogramme  $\mathcal{H}$  représentatif du nombre d'occurrences des contextes types identifiés (étape 521).

À partir des histogrammes déterminés à l'étape 513, il est alors possible de calculer une distance entre des histogrammes types (523) représentatifs de formes et/ou de mouvements, et ainsi déterminer l'histogramme type le plus proche (étape 522).

Grâce à cet histogramme type déterminé, il est possible de déterminer la forme et/ou le mouvement associé à celui-ci (étape 524) et ainsi retourner cette forme et/ou

ce mouvement (525).

Bien entendu, si l'ordinogramme de la figure 5 montre l'exécution des étapes 520 à 524 après l'analyse d'au moins un deuxième niveau hiérarchique de contextes types, ces étapes 520 à 524 peuvent tout à fait être exécutées après les étapes 501 à 505 (i.e. avec l'analyse d'un seul niveau hiérarchique).

La figure 7 illustre un dispositif pour la mise en œuvre d'un mode de réalisation selon l'invention.

Dans ce mode de réalisation, le dispositif comporte un ordinateur 700, comprenant une mémoire 705 pour stocker des instructions permettant la mise en œuvre du procédé, les données relatives au flux d'événements reçu, et des données temporaires pour réaliser les différentes étapes du procédé tel que décrit précédemment.

L'ordinateur comporte en outre un circuit 704. Ce circuit peut être, par exemple :

- un processeur apte à interpréter des instructions sous la forme de programme informatique, ou
- une carte électronique dont les étapes du procédé de l'invention sont décrites dans le silicium, ou encore
- une puce électronique programmable comme une puce FPGA (pour « Field-Programmable Gate Array » en anglais).

Cet ordinateur comporte une interface d'entrée 703 pour la réception des événements des capteurs, et une interface de sortie 706 pour la fourniture des formes 707 identifiées dans le flux d'événement. Enfin, l'ordinateur peut comporter, pour permettre une interaction aisée avec un utilisateur, un écran 701 et un clavier 702. Bien entendu, le clavier est facultatif, notamment dans le cadre d'un ordinateur ayant la forme d'une tablette tactile, par exemple.

La figure 8a est un exemple de performance de détection de formes (trèfle,



carreau, cœur et pique).

Chaque ligne (voir ordonnée 801) correspond aux données relatives à une forme particulière ayant été apprise (i.e. histogramme type, voir supra).

5 Chaque colonne (voir abscisse 802) correspond aux données relatives à une forme particulière devant être reconnues (i.e. histogramme à reconnaître).

L'intersection d'une ligne et d'une colonne (correspondant à l'ensemble 803, par exemple) permet de visualiser la distance de l'histogramme type avec l'histogramme à reconnaître pour plusieurs présentations de la forme (ici, 9 présentations de la forme, ces présentations sont séparées par un trait vertical pointillé) :

- 10
- la barre blanche (barre de gauche) correspond à une distance classique ;
  - la barre grise (barre du milieu) correspond à une distance normalisée ;
  - la barre noire (barre de droite) correspond à une distance de *Bhattacharyya*.

15 Les trois distances (respectivement classique, normalisée, et de *Bhattacharyya*) ainsi représentées possèdent une performance respective de 94%, 100% et 97% pour ces formes.

La figure 8b est un exemple de performance de détection de formes (chiffres et lettres majuscules).

20 Chaque ligne (voir ordonnée 804) correspond aux données relatives à une forme particulière ayant été apprise (i.e. histogramme type, voir supra).

Chaque colonne (voir abscisse 805) correspond aux données relatives à une forme particulière devant être reconnues (i.e. histogramme à reconnaître).

25 L'intersection d'une ligne et d'une colonne (correspondant à l'ensemble 806, par exemple) permet de visualiser la distance de l'histogramme type avec l'histogramme à reconnaître pour une présentation de la forme :

- la barre blanche (barre de gauche) correspond à une distance classique ;
- la barre grise (barre du milieu) correspond à une distance normalisée ;

- la barre noire (barre de droite) correspond à une distance de *Bhattacharyya*.

Les trois distances (respectivement classique, normalisée, et de *Bhattacharyya*) ainsi représentées possèdent toutes une performance de 100% pour ces formes.

5 La figure 8c est un exemple de performance de détection de formes (visages).

Chaque tableau (807, 808 et 809) présente le nombre d'occurrences de reconnaissance d'un visage présenté (abscisse) à l'aide de visage appris (ordonnées) pour 19 présentations de ce visage.

10 Le tableau 807 utilise une distance classique pour effectuer la reconnaissance de visage (taux de reconnaissance : 37%).

Le tableau 808 utilise une distance normalisée pour effectuer la reconnaissance de visage (taux de reconnaissance : 78%).

Le tableau 808 utilise une distance de *Bhattacharyya* pour effectuer la reconnaissance de visage (taux de reconnaissance : 79%).

15

Par ailleurs, le schéma fonctionnel présenté sur la figure 5 est un exemple typique d'un programme dont certaines instructions peuvent être réalisées auprès du dispositif décrit. À ce titre, la figure 5 peut correspondre à l'organigramme de l'algorithme général d'un programme informatique au sens de l'invention.

20

Bien entendu, la présente invention ne se limite pas aux formes de réalisation décrites ci-avant à titre d'exemple ; elle s'étend à d'autres variantes.

D'autres réalisations sont possibles.

25 Par exemple, les contextes types des figures 6a à 6c peuvent être de n'importe quelles dimensions.

Par ailleurs, la description mentionne principalement des sources d'événements en provenance d'un capteur de lumière/vidéo. Néanmoins, l'invention décrite se

généralise à tout ensemble de sources d'évènements comme par exemple un réseau de capteurs à pression qui fonctionneraient sur un tel mode asynchrone et dont l'agencement spatial serait assimilable à un capteur vidéo.

## **REVENDEICATIONS**

1. Procédé de traitement d'un signal asynchrone produit par un capteur de lumière, le capteur ayant une matrice de pixels disposée en regard d'une scène, le  
5 procédé comprenant :
- recevoir du capteur de lumière le signal asynchrone (500) comprenant, pour chaque pixel de la matrice, des événements successifs issus dudit pixel ;
  - analyser (501) le signal asynchrone à l'aide de profils d'activité au fur et à mesure de la réception d'événements dans le signal asynchrone ;
- 10 dans lequel le profil d'activité comprend au moins, pour chaque pixel ( $p$ ) du capteur, une valeur d'activité ( $S(p)$ ) qui décroît en fonction du temps ( $t$ ) écoulé depuis un événement le plus récent (310, 311, 312, 313, 314, 320, 321, 322) parmi les événements successifs issus dudit pixel.
- 15 2. Procédé selon la revendication 1, dans lequel le profil d'activité décroît exponentiellement en fonction du temps ( $t$ ) écoulé depuis un événement le plus récent (310, 311, 312, 313, 314, 320, 321, 322) parmi les événements successifs issus dudit pixel.
- 20 3. Procédé selon l'une des revendications précédentes, dans lequel le profil d'activité est en outre en fonction du temps ( $t$ ) écoulé depuis un événement antérieur à l'événement le plus récent (310, 311, 312, 313, 314, 320, 321, 322) parmi les événements successifs issus dudit pixel.
- 25 4. Procédé de reconnaissance de formes, le procédé comprend :
- /a/ l'exécution du procédé de traitement d'un signal asynchrone selon l'une des revendications précédentes ;
  - /b/ au fur et à mesure de la réception des événements, pour chaque pixel ( $p$ )

courant dudit capteur :

- 5            /b1/ détermination (501) d'un contexte ( $S_p$ ) pour ledit pixel courant, ledit contexte ( $S_p$ ) étant défini comme un ensemble de profils d'activité de pixels situés à moins d'une distance prédéterminée dudit pixel ( $p$ ) courant,
- /b2/ détermination d'un contexte type ( $C_k$ ) parmi un ensemble de contextes types (504) prédéterminés, le contexte type déterminé ayant, parmi l'ensemble des contextes types, une distance au contexte déterminé à l'étape /b1/ minimum,
- 10           /c/ détermination (521) d'une signature fonction d'un nombre d'occurrences de détermination des contextes types à l'étape /b2/ ;
- /d/ détermination (524) d'une forme par une comparaison de ladite signature avec une base de signature types.
- 15           5. Procédé de reconnaissance de formes, le procédé comprend, un modèle hiérarchique de contextes types étant définis, chaque contexte type étant associé à une pluralité de niveau du modèle hiérarchique :
- /a/ utilisation d'événements du signal asynchrone du capteur comme événements courants et utilisation des contextes types du premier niveau du
- 20           modèle hiérarchique comme contextes types courants ;
- /b/ exécution du procédé de traitement d'un signal asynchrone selon l'une des revendications 1 à 3 avec les événements courants ;
- /c/ au fur et à mesure de la réception des événements courants, pour chaque pixel ( $p$ ) courant dudit capteur :
- 25           /c1/ détermination (501) d'un contexte ( $S_p$ ) pour ledit pixel courant, ledit contexte ( $S_p$ ) étant défini comme un ensemble de profils d'activité utilisés à l'étape /b/ pour des pixels situés à moins d'une distance prédéterminée dudit pixel ( $p$ ) courant,

/c2/ détermination d'un contexte type ( $C_k$ ) parmi les contextes types courants (504), le contexte type déterminé ayant, parmi l'ensemble des contextes types courants, une distance au contexte déterminé à l'étape /c1/ minimum,

5            /c3/ génération d'un événement fonction du contexte type identifié à l'étape /c2/ ;

/d/ si un niveau du modèle hiérarchique n'a pas été utilisé :

- utilisation des événements générés à l'étape /c3/ comme événements courants,

10           - utilisation des contextes types d'un niveau du modèle hiérarchique non encore utilisé comme contextes types courants ;

- réexécution des étapes /b/ à /e/ ;

/e/ si tous les niveaux du modèle hiérarchique ont été utilisés :

15           - détermination (521) d'une signature fonction d'un nombre d'occurrences de détermination des contextes types à la dernière occurrence de l'étape /c2/ ;

- détermination (524) d'une forme par une comparaison de ladite signature avec une base de signatures types.

20           6. Procédé selon l'une des revendications 4 à 5, dans lequel la détermination d'un contexte prend en compte, séparément, les événements ayant des polarités différentes.

25           7. Procédé selon l'une des revendications 5 et 6, dans lequel la comparaison de l'étape /e/ comporte le calcul d'une distance de *Bhattacharyya*.

8. Procédé selon l'une des revendications 5 et 6, dans lequel la comparaison de

l'étape /e/ comporte le calcul d'une distance normalisée.

9. Produit-programme informatique comportant des instructions pour la mise en œuvre du procédé selon l'une des revendications 1 à 8, lorsque ce programme est
- 5 exécuté par un processeur.

1/13

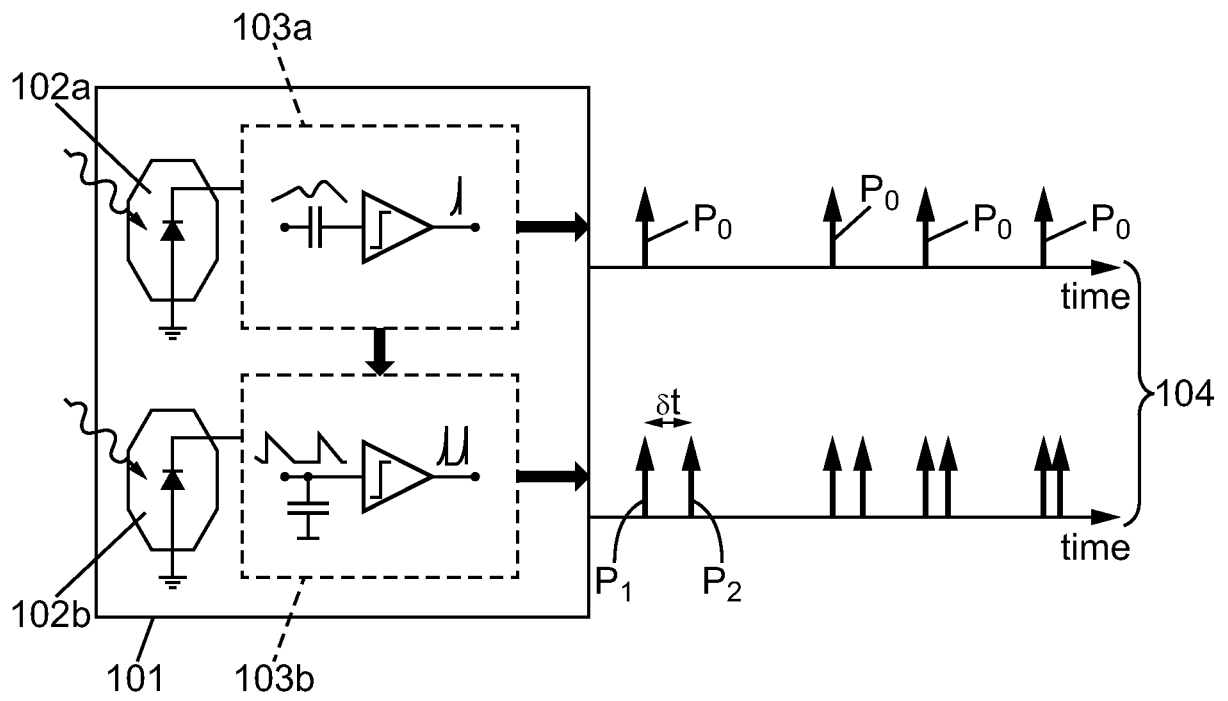


FIG. 1

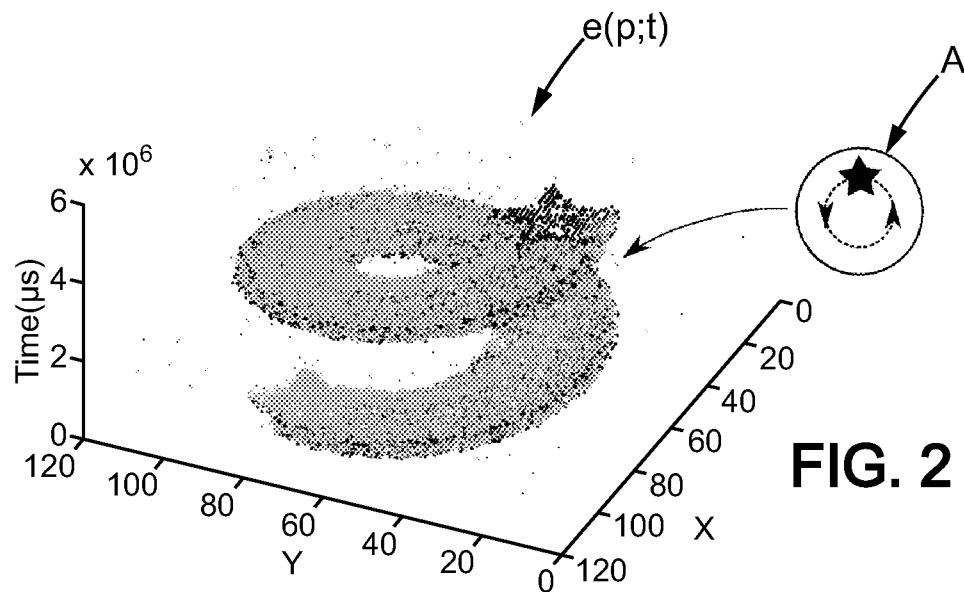


FIG. 2



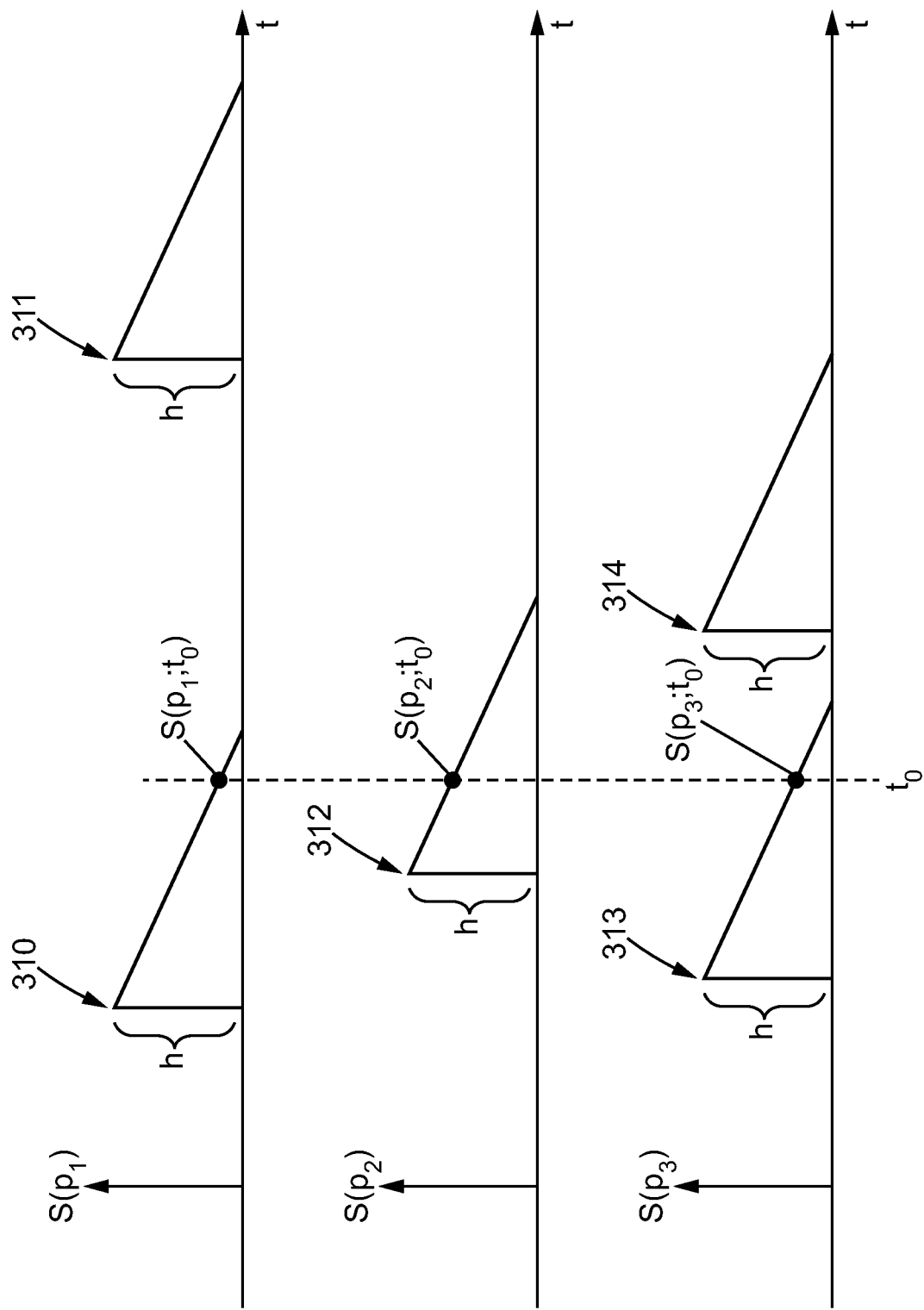


FIG. 3a

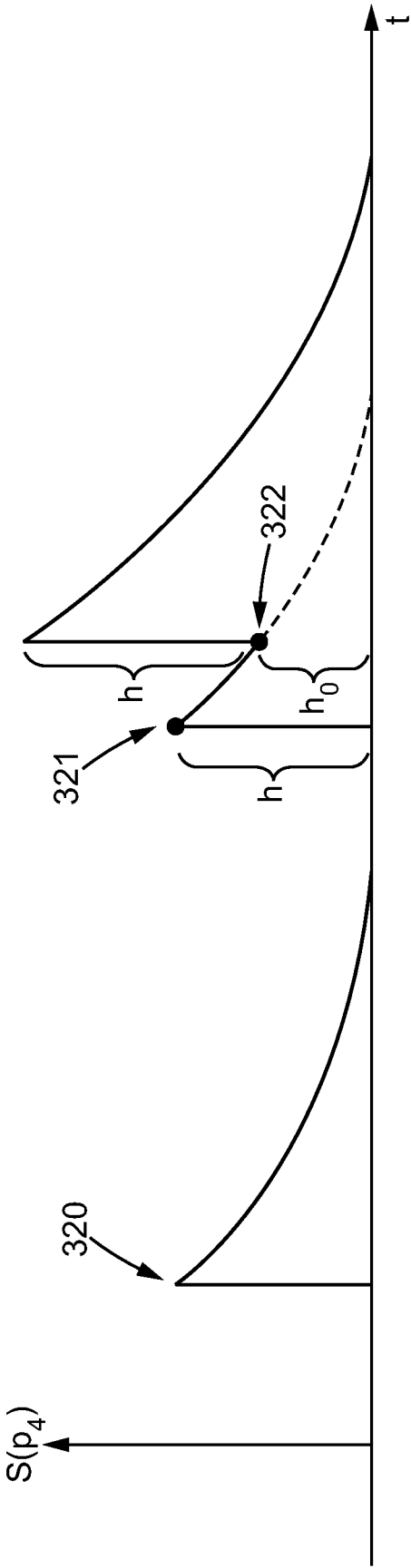


FIG. 3b

4/13

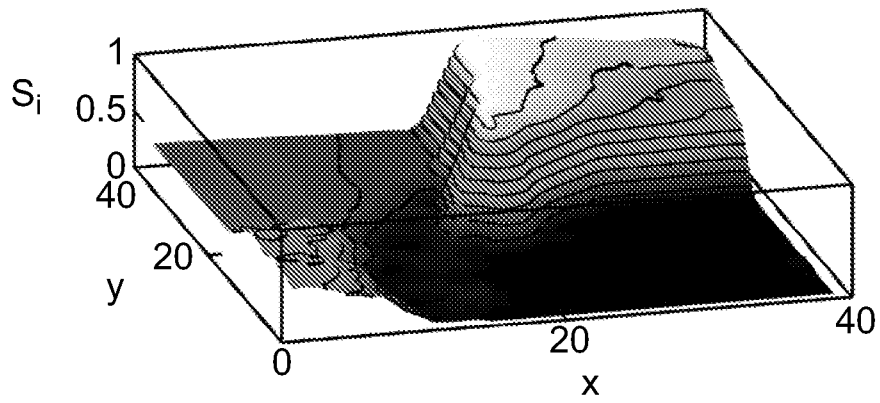


FIG. 3c

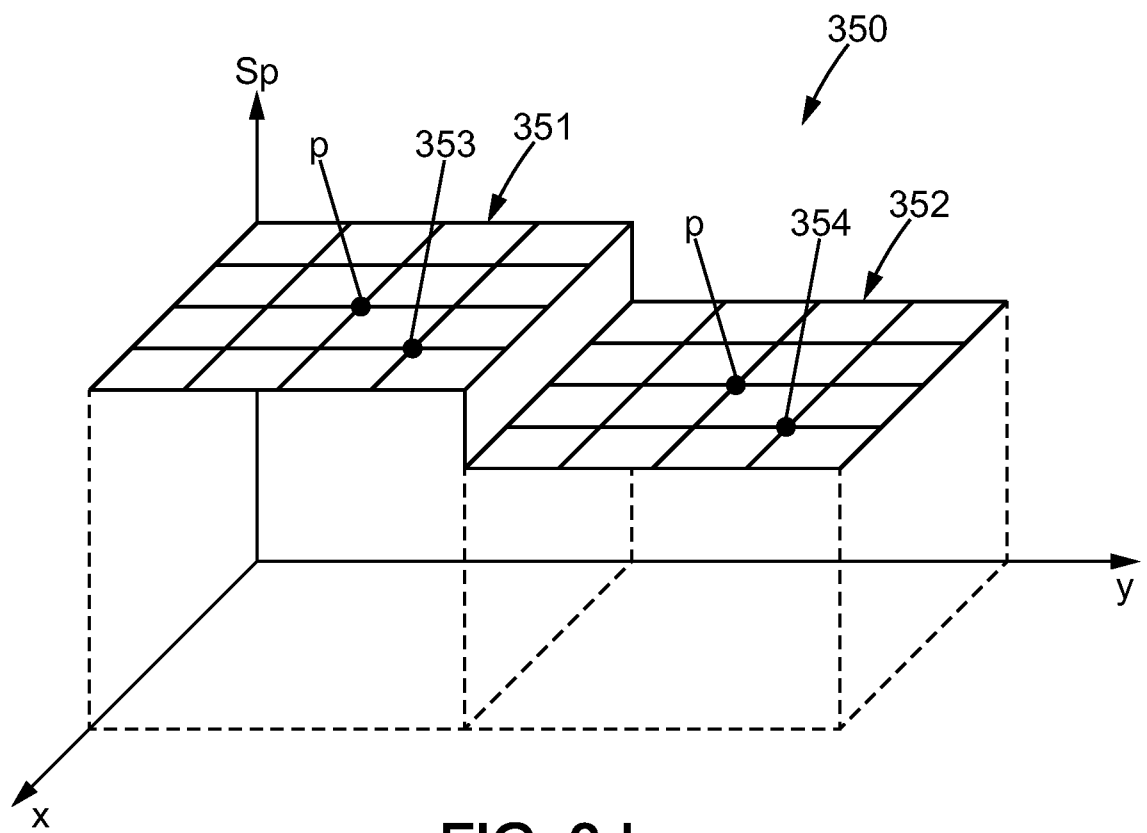


FIG. 3d

5/13

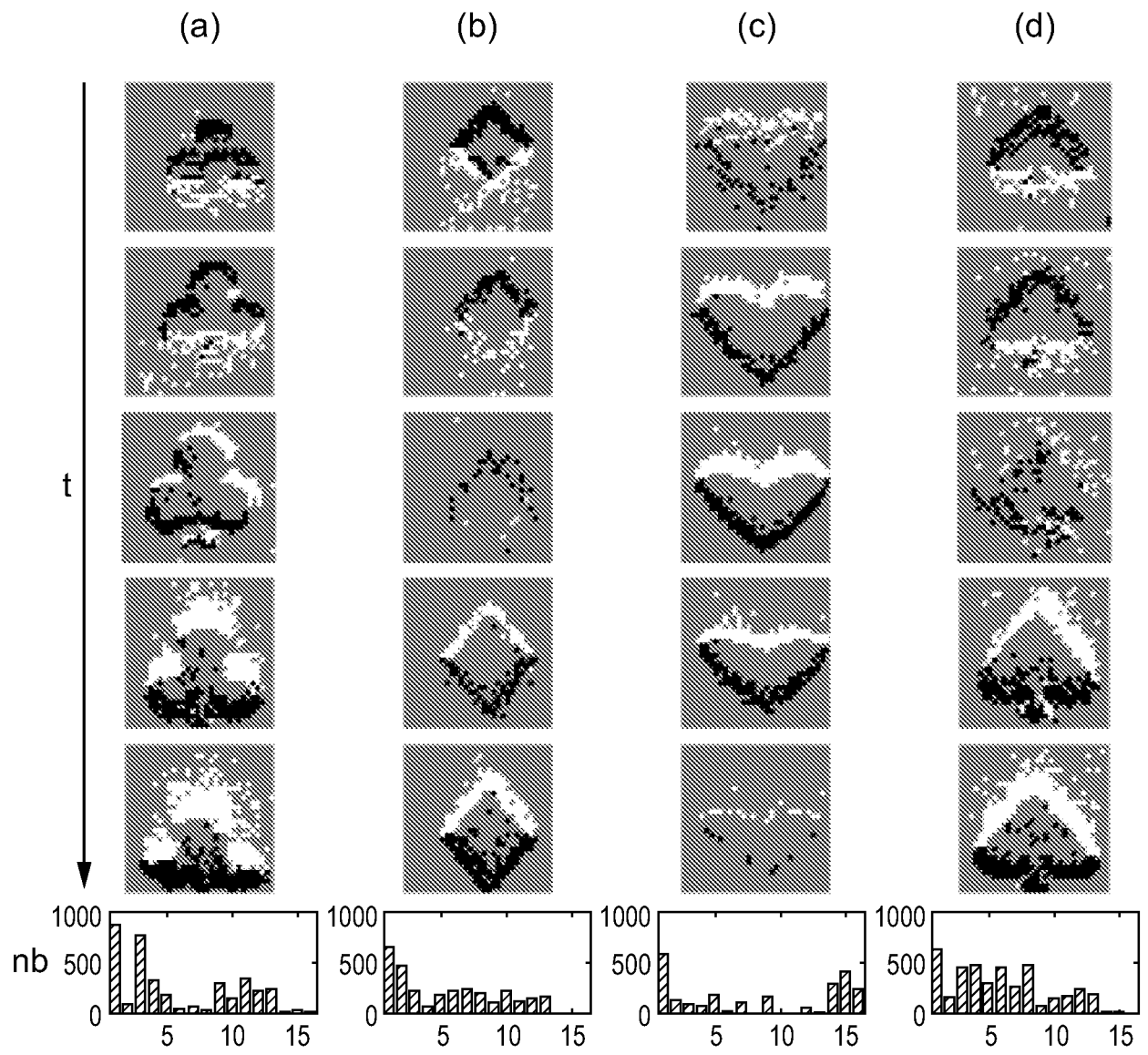
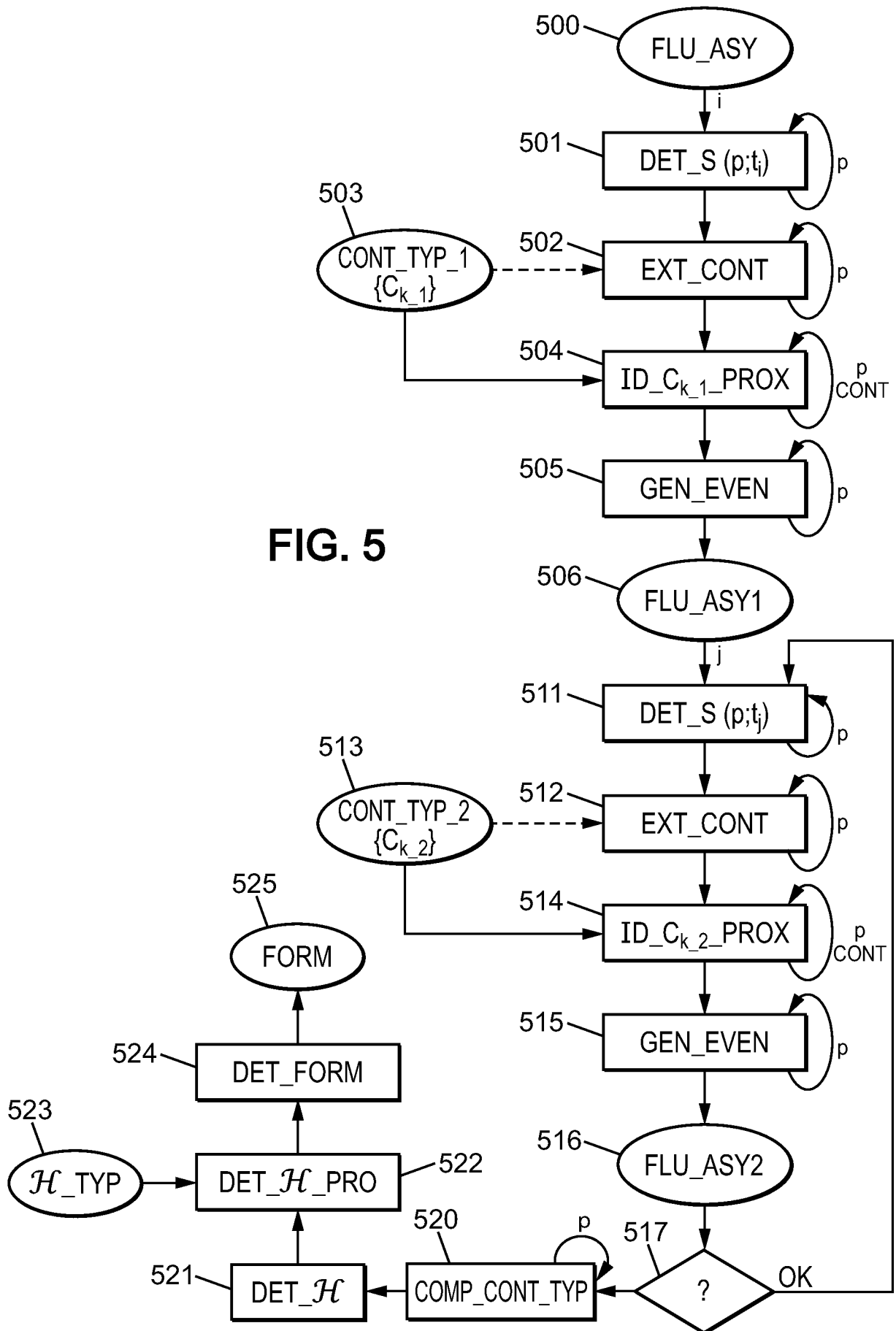


FIG. 4

6/13



7/13

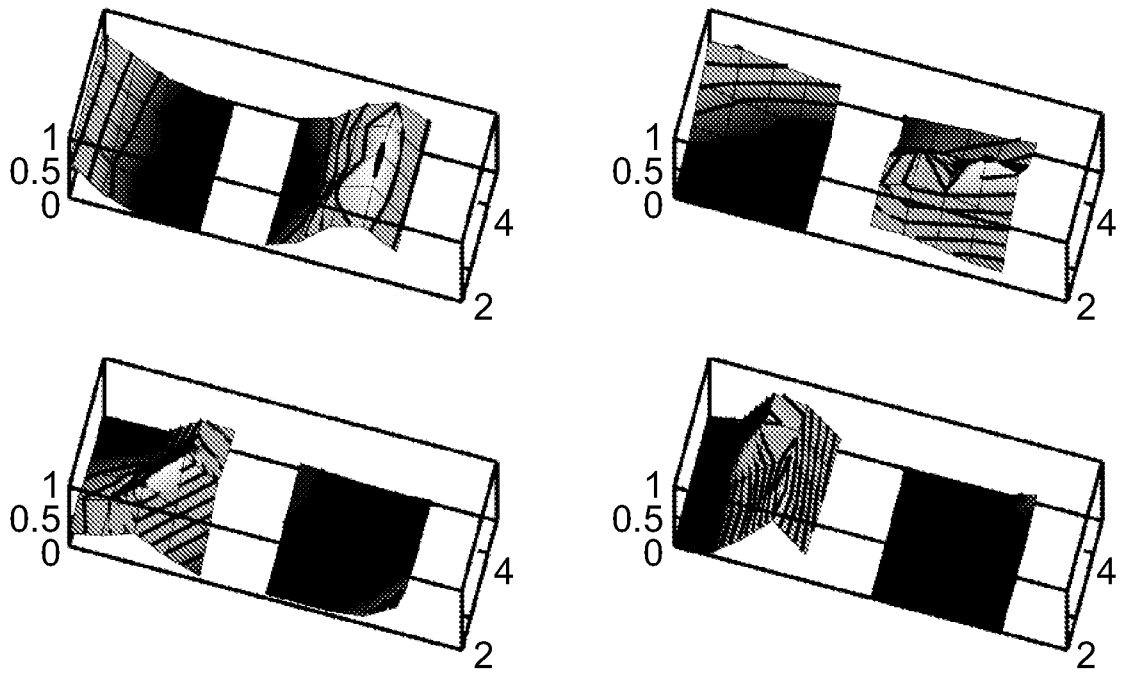


FIG. 6a

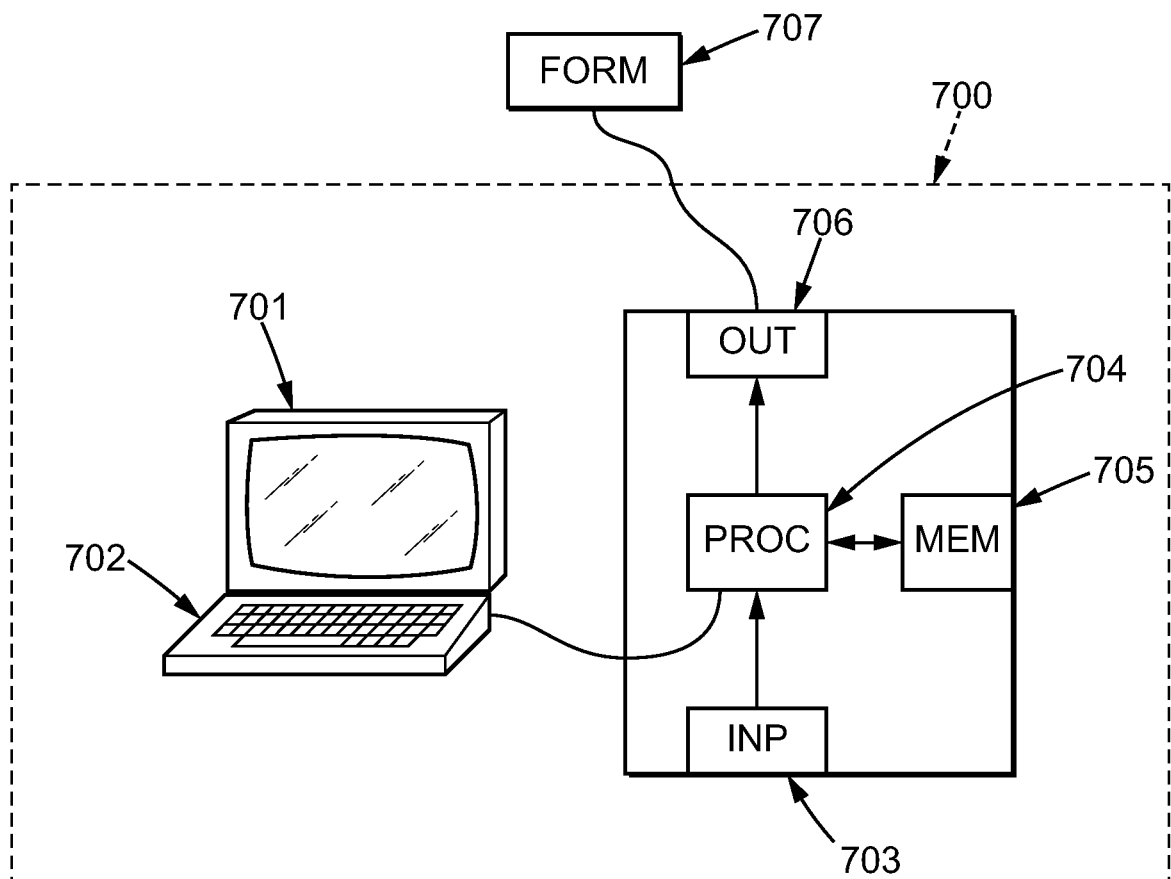


FIG. 7

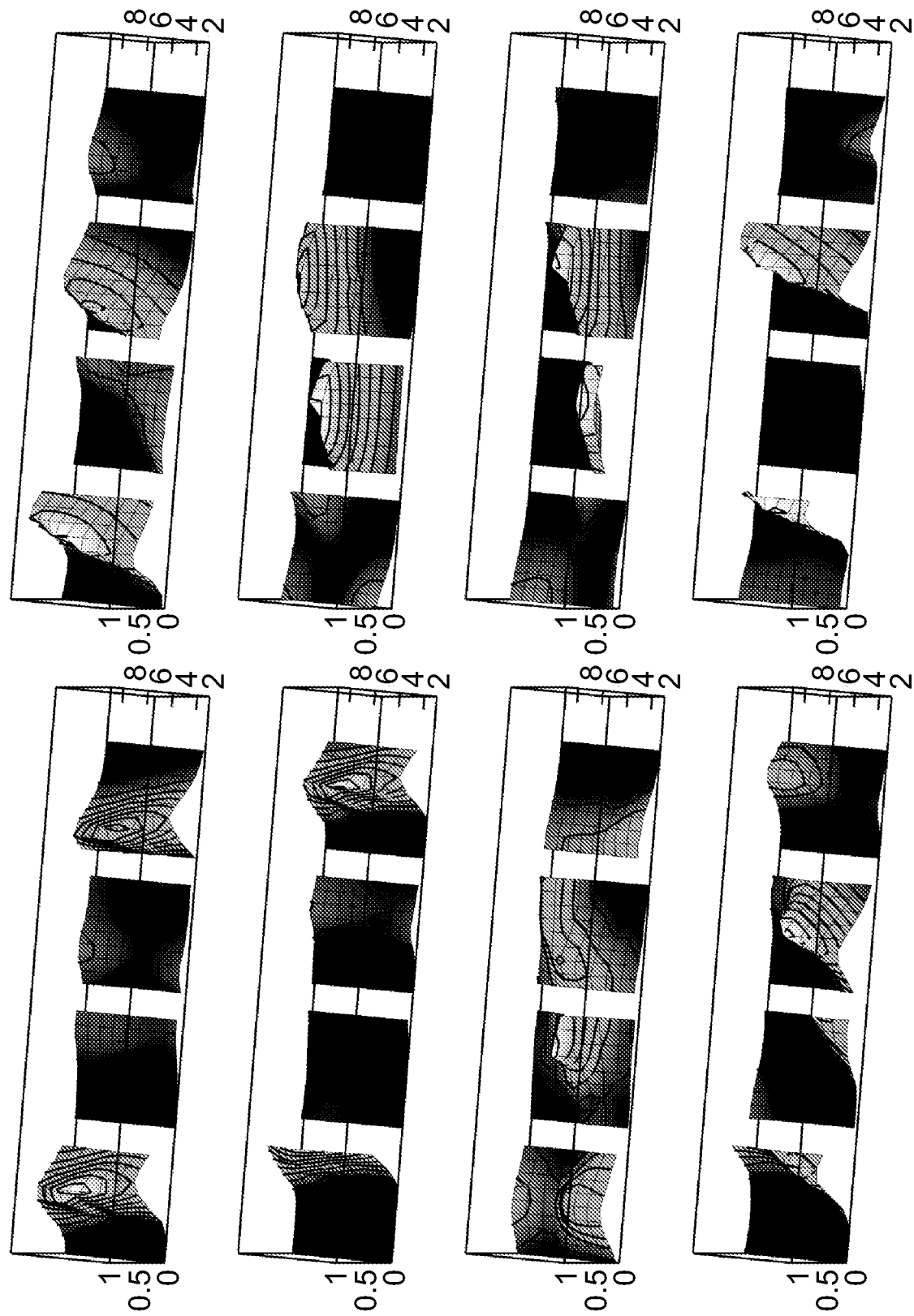


FIG. 6b

9/13

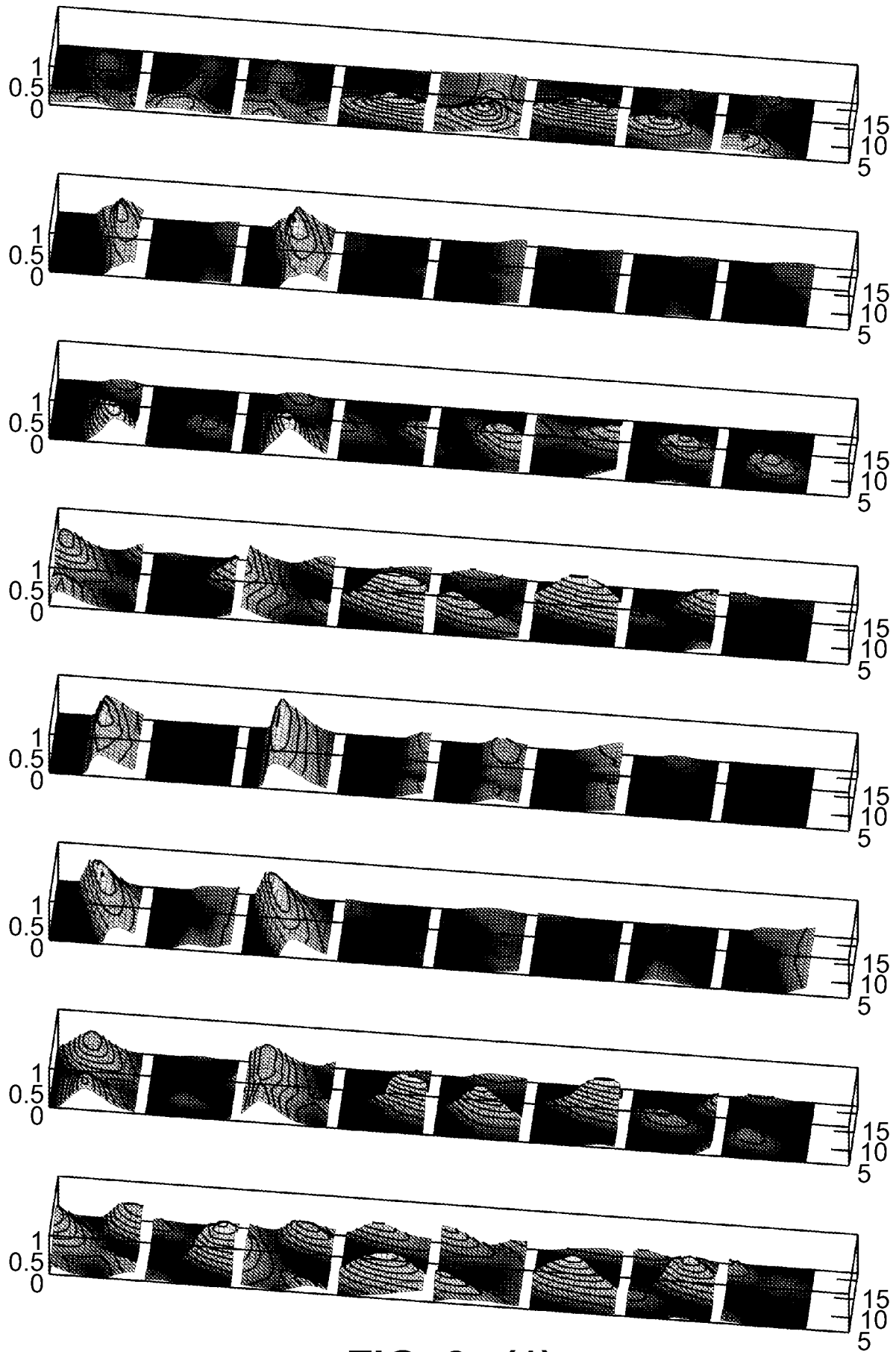


FIG. 6c (1)



10/13

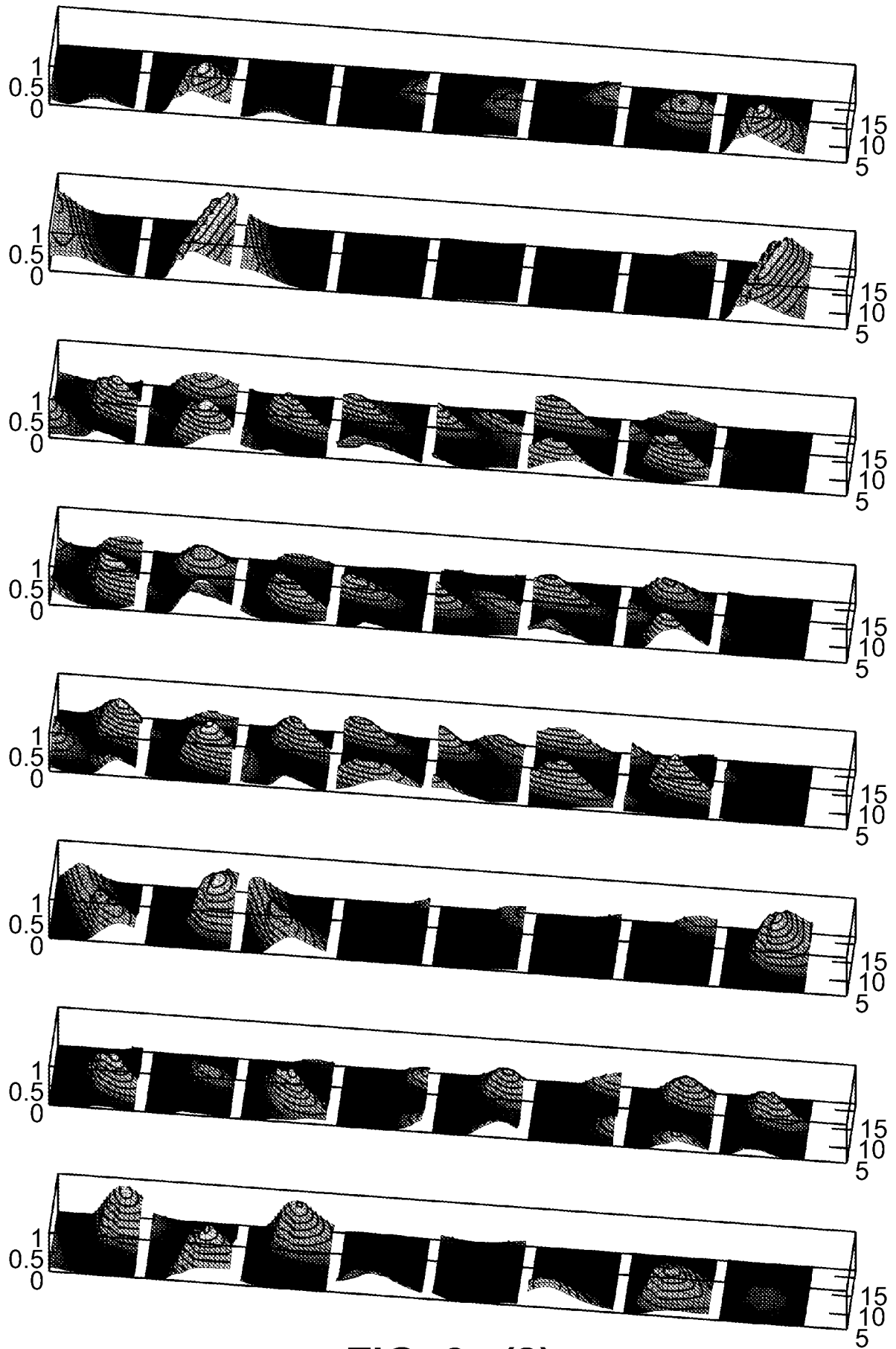


FIG. 6c (2)

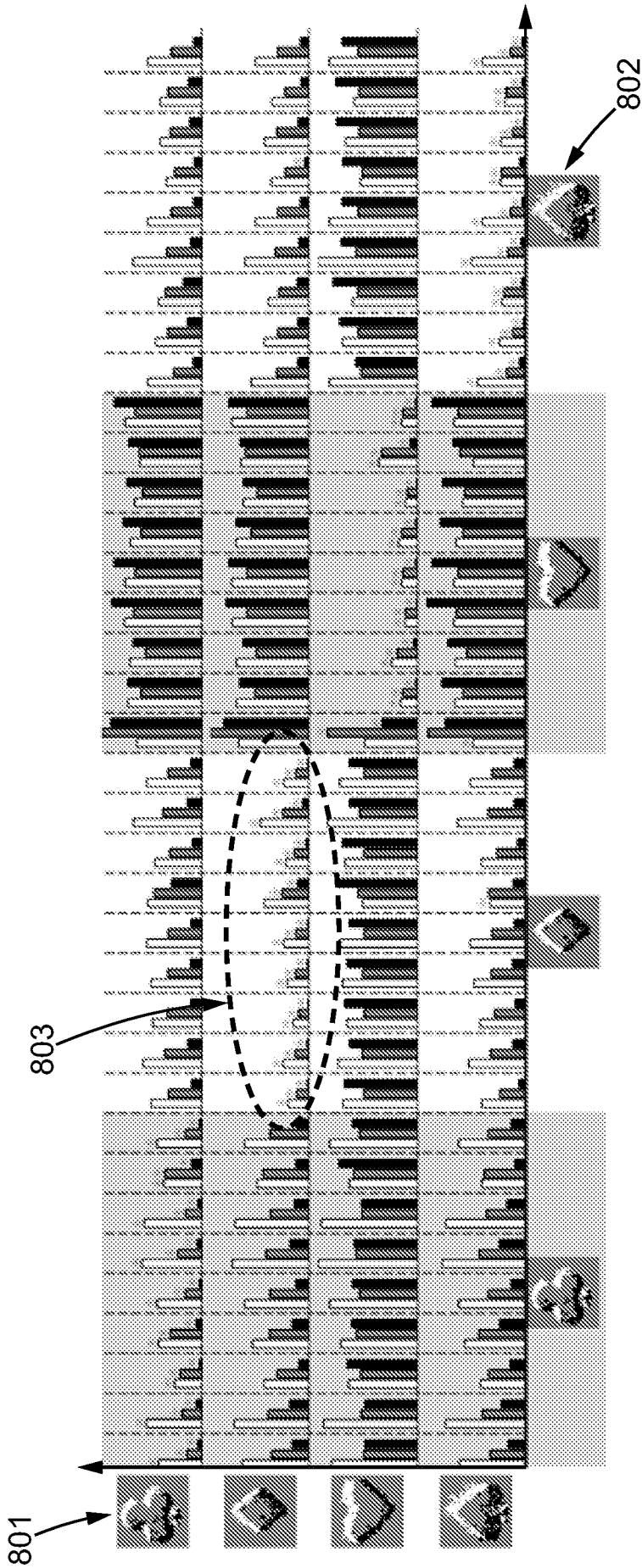
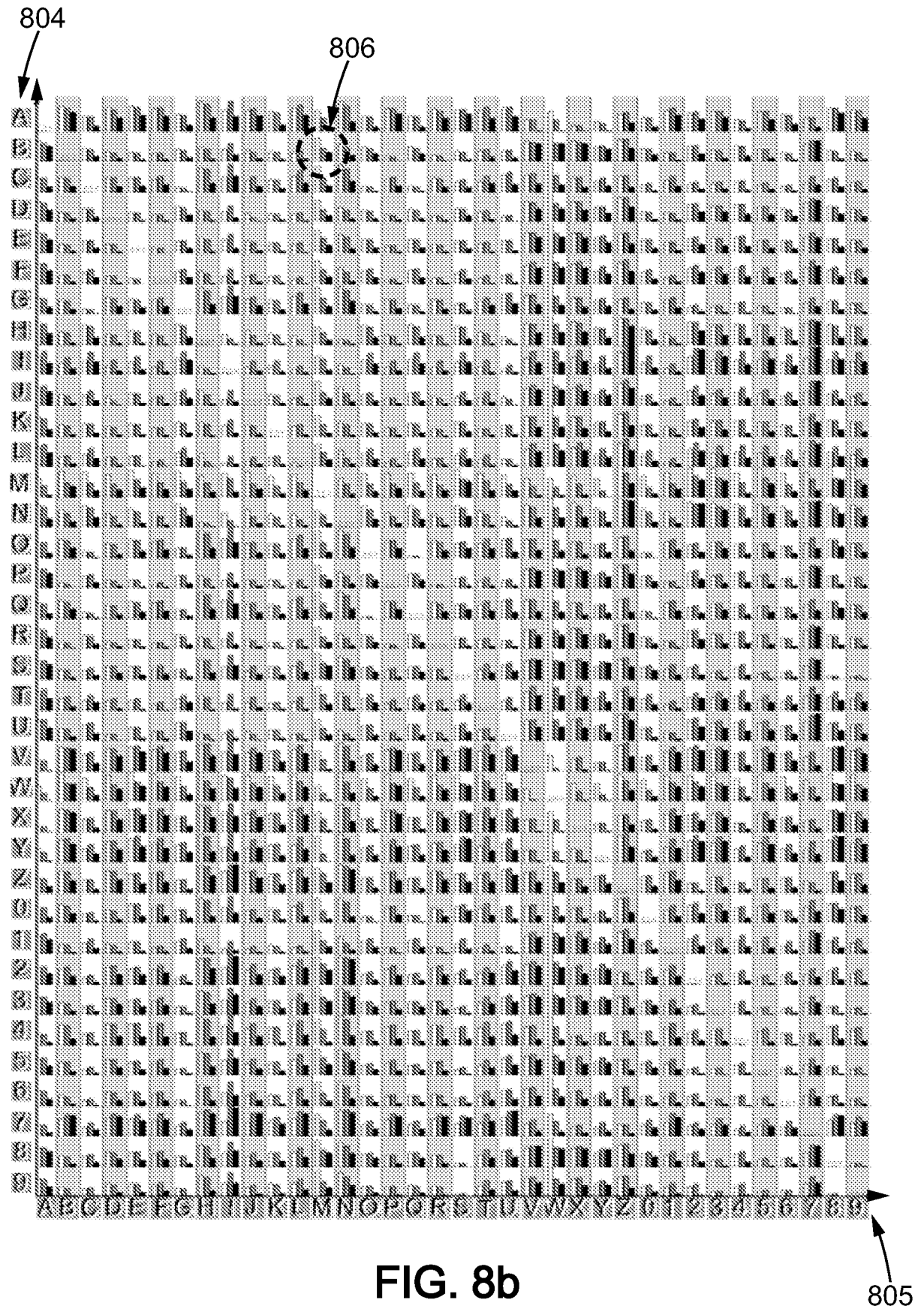


FIG. 8a

12/13



13/13

807 →

(a)							
	11		1	16			3
		5					3
		1					1
	6	6	9	3	1		5
	2	5	9		10		6
						19	
		2			8		1

← 808

(b)							
	14	2	4				
		15	2	1	1		1
			9	2			
	3			14			
			1		16		
	2	2	1	2	2	19	1
			2				17

809 →

(c)							
	14						
		17			2		2
		1	13	4			
	3		4	13	1		
					16		4
	2	1	2	2		19	
							13

FIG. 8c

## INTERNATIONAL SEARCH REPORT

International application No

PCT/FR2016/050574

## A. CLASSIFICATION OF SUBJECT MATTER

INV. G06K9/62

ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal, WPI Data

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	ZHENJIANG NI ET AL: "Asynchronous Event-Based Visual Shape Tracking for Stable Haptic Feedback in Microrobotics", IEEE TRANSACTIONS ON ROBOTICS, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 28, no. 5, 1 October 2012 (2012-10-01), pages 1081-1089, XP011474264, ISSN: 1552-3098, DOI: 10.1109/TRO.2012.2198930	1-3,9
A	section III.A ----- -/--	4-8



Further documents are listed in the continuation of Box C.



See patent family annex.

## \* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&amp;" document member of the same patent family

Date of the actual completion of the international search

13 June 2016

Date of mailing of the international search report

20/06/2016

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Authorized officer

Martinez, Francis

## INTERNATIONAL SEARCH REPORT

International application No  
PCT/FR2016/050574

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GEORG WIESMANN ET AL: "Event-driven embodied system for feature extraction and object recognition in robotic applications", COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), 2012 IEEE COMPUTER SOCIETY CONFERENCE ON, IEEE, 16 June 2012 (2012-06-16), pages 76-82, XP032206761, DOI: 10.1109/CVPRW.2012.6238898 ISBN: 978-1-4673-1611-8	1-3,9
A	abstract figure 3 section 2 - section 3.3 -----	4-8
X	SCHRAML S ET AL: "A real-time pedestrian classification method for event-based dynamic stereo vision", COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), 2010 IEEE COMPUTER SOCIETY CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 13 June 2010 (2010-06-13), pages 93-99, XP031728957, ISBN: 978-1-4244-7029-7	1-3,9
A	abstract section 2 - section 3.2 -----	4-8
X	GARRICK ORCHARD ET AL: "HFirst: A Temporal Approach to Object Recognition", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 37, no. 10, 14 January 2015 (2015-01-14), pages 2028-2040, XP055242818, USA ISSN: 0162-8828, DOI: 10.1109/TPAMI.2015.2392947	1-3,9
A	abstract section 2 - section 4.3 -----	4-8

# RAPPORT DE RECHERCHE INTERNATIONALE

Demande internationale n°

PCT/FR2016/050574

A. CLASSEMENT DE L'OBJET DE LA DEMANDE  
INV. G06K9/62  
ADD.

Selon la classification internationale des brevets (CIB) ou à la fois selon la classification nationale et la CIB

## B. DOMAINES SUR LESQUELS LA RECHERCHE A PORTE

Documentation minimale consultée (système de classification suivi des symboles de classement)  
G06K

Documentation consultée autre que la documentation minimale dans la mesure où ces documents relèvent des domaines sur lesquels a porté la recherche

Base de données électronique consultée au cours de la recherche internationale (nom de la base de données, et si cela est réalisable, termes de recherche utilisés)  
EPO-Internal, WPI Data

## C. DOCUMENTS CONSIDERES COMME PERTINENTS

Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	ZHENJIANG NI ET AL: "Asynchronous Event-Based Visual Shape Tracking for Stable Haptic Feedback in Microrobotics", IEEE TRANSACTIONS ON ROBOTICS, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 28, no. 5, 1 octobre 2012 (2012-10-01), pages 1081-1089, XP011474264, ISSN: 1552-3098, DOI: 10.1109/TRO.2012.2198930	1-3,9
A	section III.A ----- -/-	4-8



Voir la suite du cadre C pour la fin de la liste des documents



Les documents de familles de brevets sont indiqués en annexe

\* Catégories spéciales de documents cités:

"A" document définissant l'état général de la technique, non considéré comme particulièrement pertinent

"E" document antérieur, mais publié à la date de dépôt international ou après cette date

"L" document pouvant jeter un doute sur une revendication de priorité ou cité pour déterminer la date de publication d'une autre citation ou pour une raison spéciale (telle qu'indiquée)

"O" document se référant à une divulgation orale, à un usage, à une exposition ou tous autres moyens

"P" document publié avant la date de dépôt international, mais postérieurement à la date de priorité revendiquée

"T" document ultérieur publié après la date de dépôt international ou la date de priorité et n'appartenant pas à l'état de la technique pertinent, mais cité pour comprendre le principe ou la théorie constituant la base de l'invention

"X" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme nouvelle ou comme impliquant une activité inventive par rapport au document considéré isolément

"Y" document particulièrement pertinent; l'invention revendiquée ne peut être considérée comme impliquant une activité inventive lorsque le document est associé à un ou plusieurs autres documents de même nature, cette combinaison étant évidente pour une personne du métier

"&" document qui fait partie de la même famille de brevets

Date à laquelle la recherche internationale a été effectivement achevée

13 juin 2016

Date d'expédition du présent rapport de recherche internationale

20/06/2016

Nom et adresse postale de l'administration chargée de la recherche internationale

Office Européen des Brevets, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040,  
Fax: (+31-70) 340-3016

Fonctionnaire autorisé

Martinez, Francis

C(suite). DOCUMENTS CONSIDERES COMME PERTINENTS		
Catégorie*	Identification des documents cités, avec, le cas échéant, l'indication des passages pertinents	no. des revendications visées
X	GEORG WIESMANN ET AL: "Event-driven embodied system for feature extraction and object recognition in robotic applications", COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), 2012 IEEE COMPUTER SOCIETY CONFERENCE ON, IEEE, 16 juin 2012 (2012-06-16), pages 76-82, XP032206761, DOI: 10.1109/CVPRW.2012.6238898 ISBN: 978-1-4673-1611-8	1-3,9
A	abrégé figure 3 section 2 - section 3.3 -----	4-8
X	SCHRAML S ET AL: "A real-time pedestrian classification method for event-based dynamic stereo vision", COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), 2010 IEEE COMPUTER SOCIETY CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 13 juin 2010 (2010-06-13), pages 93-99, XP031728957, ISBN: 978-1-4244-7029-7	1-3,9
A	abrégé section 2 - section 3.2 -----	4-8
X	GARRICK ORCHARD ET AL: "HFirst: A Temporal Approach to Object Recognition", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, vol. 37, no. 10, 14 janvier 2015 (2015-01-14), pages 2028-2040, XP055242818, USA ISSN: 0162-8828, DOI: 10.1109/TPAMI.2015.2392947	1-3,9
A	abrégé section 2 - section 4.3 -----	4-8





## (12)发明专利申请

(10)申请公布号 CN 107873097 A

(43)申请公布日 2018.04.03

(21)申请号 201680019785.4

(22)申请日 2016.03.15

(30)优先权数据

1552155 2015.03.16 FR

(85)PCT国际申请进入国家阶段日

2017.09.29

(86)PCT国际申请的申请数据

PCT/FR2016/050574 2016.03.15

(87)PCT国际申请的公布数据

W02016/146937 FR 2016.09.22

(71)申请人 皮埃尔和玛利居里大学(巴黎第六大学)

地址 法国巴黎

申请人 国家科学研究中心

法国国家健康和医学研究院

(72)发明人 伊恩格·西奥-霍伊

贝诺斯梅恩·尔亚德

施·伯特拉姆

(74)专利代理机构 上海天协和诚知识产权代理  
事务所 31216

代理人 张轶

(51)Int.Cl.

G06K 9/62(2006.01)

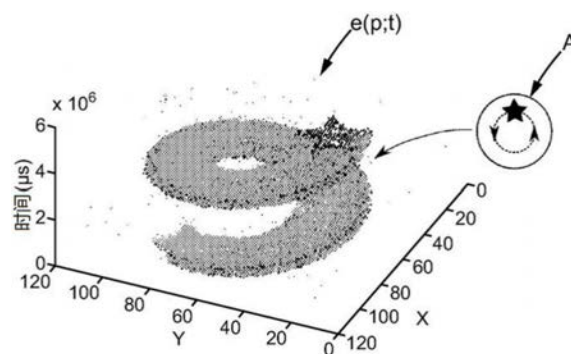
权利要求书2页 说明书12页 附图13页

(54)发明名称

适用于处理异步信号的方法

(57)摘要

本发明涉及处理由光传感器所产生的异步信号的方法,所述传感器具有施置面对场景的像素矩阵,所述方法包括:-接收光传感器的异步信号(500),相对于矩阵的每个像素而言,所述异步信号包括来自所述像素的连续事件;-当接收到异步信号中的事件,利用活动简档来分析异步信号(501)。活动简档至少包括针对于传感器的每个像素(p)的活动数值( $S(p)$ ),所述活动数值随着从来自所述像素的连续事件中最近的事件(310、311、312、313、314、320、321、322)起始之后所经过的时间(t)减少。



1. 适用于处理异步信号的方法,所述异步信号是由光传感器产生的,所述传感器的像素矩阵位于面对场景的位置,所述方法包括:

-接收光传感器的异步信号(500),对于矩阵的每个像素而言,所述异步信号包括来自所述像素的连续事件;

-当在异步信号中接收到事件,利用活动简档来分析异步信号(501);

其中,对于传感器的每个像素(p)而言,活动简档至少包括活动数值(s(p)),所述活动数值随着从来自所述像素的连续事件中最近的事件(310、311、312、313、314、320、321、322)起始之后所经过的时间(t)减少。

2. 根据权利要求1所述的方法,其特征在于,所述活动简档根据从来自所述像素的连续事件中最近的事件(310、311、312、313、314、320、321、322)起始之后所经过的时间(t)以指数方式减少。

3. 根据前述权利要求中任一项所述的方法,其特征在于,所述活动简档进一步为时间(t)函数,所述时间是来自所述像素的连续事件中最近的事件(310、311、312、313、314、320、321、322)之前的事件起始之后所经过的时间。

4. 适用于识别形状的方法,所述方法包括:

/a/执行根据前述权利要求中任一项所述的处理异步信号的方法;

/b/当接收到事件时,针对所述传感器的每个当前像素(p):

/b1/确定(501)所述当前像素的背景( $S_p$ ),所述背景( $S_p$ )定义为位置距离与所述当前像素(p)小于预定距离的像素的一组活动简档,

/b2/从一组预定典型背景(504)中确定典型背景( $C_k$ ),使从一组典型背景中所确定的典型背景与在步骤/b1/中所确定背景的距离最小,

/c/根据步骤/b2/中所确定的典型背景的出现次数,来确定(521)识别标志;

/d/通过比较所述识别标志与典型识别标志类型的基础来确定(524)形状。

5. 适用于识别形状的方法,所述方法包括:确定典型背景的分层模型,将每个典型背景都与分层模型的多个等级相关联:

/a/将传感器异步信号的事件用作当前事件,把分层模型的第一级别的典型背景用作当前典型背景;

/b/执行根据权利要求1至3中任一项所述的针对当前事件所述的处理异步信号的方法;

/c/当接收到当前事件时,对于所述传感器的每个当前像素(p)而言:

/c1/确定(501)所述当前像素的背景( $S_p$ ),将所述背景( $S_p$ )定义为在步骤/b/中所用于位置距离与所述当前像素(p)相间隔的距离小于预定距离的像素的一组活动简档,

/c2/从当前典型背景(504)中确定典型背景( $C_k$ ),确定的典型背景在所有当前典型背景之中与步骤/c1/中所确定的背景的距离最小,

/c3/根据在步骤/c2/中所识别的典型背景来生成事件;

/d/如果尚未使用一个级别的分层模型:

-将在步骤/c3/中所生成的事件用作当前事件,

-将尚未使用的一个级别的分层模型的典型背景用作当前典型背景典型背景;

-再次执行步骤/b/至/e/;

/e/如果已经使用了所有级别的分层模型：

-根据在步骤/c2/中最后出现的典型背景出现的次数的确定，来确定识别标志(521)；

-通过比较所述识别标志与典型识别标志的基础来确定形状(524)。

6.根据权利要求4至5中任一项所述的方法，其特征在于，所述背景的确定分别把具有不同极性的事件考虑在内。

7.根据权利要求5至6中任一项所述的方法，其特征在于，所述步骤/e/中的比较包括巴氏距离计算。

8.根据权利要求5至6中任一项所述的方法，其特征在于，所述步骤/e/中的比较包括标准化距离的计算。

9.计算机程序产品，包括在通过处理器执行该程序的情况下执行根据权利要求1至8中任一项所述的方法的指令。

## 适用于处理异步信号的方法

### 技术领域

[0001] 本发明涉及到异步信号处理领域,尤其适用于检测异步信号的形状。

### 背景技术

[0002] 在历史上已经开发了各种视频信号处理技术。通常都是基于普通的光栅图像方法。

[0003] 现有异步传感器(DVS、ATIS)。这些传感器可有利于代替常规的摄像机。

[0004] 但是,信号处理的直观性明显不如光栅图像摄像机。尽管已经开发了一些算法用于处理一定种类的问题(例如,光流、形状跟踪),但是仍有一些其它问题疏于或者尚未探讨,或者这些问题解决起来比较棘手。

[0005] 对于形状识别而言尤其是这样。

[0006] 为了检测形状和/或运动,目前的方法(即,利用来自常规摄像机的视频数据)试图识别适合位于所讨论的视频数据附近一组像素的某些视觉特征。

[0007] 本领域技术人员大多将这些视觉特征理解为图像的空间信息(尽管该空间信息可处于运动中)。

[0008] 因此,视频数据的时间信息常常是可以忽略的:最多,可以在视频数据的两个或多个图像(帧)之间找到视觉特征的变化。

[0009] 首先可以通过通常的视频采集技术来解释时间分量的忽略:视频数据是产生大量静态图像(或帧)的采集结果。

[0010] 在历史上该视频采集技术已经调整了处理或显示视频的方式。这一系列静态图像的存在使其难以控制视频数据的时间(或动态)信息。

[0011] 尽管有可能增加每秒视频数据的图像数量,但是超过100Hz的图像频率罕见,无论出于关于采集方面的硬件限制的原因或者出于通过形状检测的常规工具对该数据进行实时处理的原因,皆是如此。

[0012] 就此而言,需要检测可靠的形状,并且可以通过充分利用视频数据的时间信息对其加以使用。

### 发明内容

[0013] 本发明旨在改进现状。

[0014] 与以定期采样时刻来记录连续图像的常规摄像机相反,已经研发了受人眼活动启发的生物视网膜。生物视网膜仅传递关于待查看场景的很少冗余信息,而且这是不同步的。

[0015] 事件驱动型异步视觉传感器可以事件的形式来传送压缩的数字数据。

[0016] T.Delbrück等人编写的《活动驱动型基于事件的视觉传感器(Activity-Driven, Event-Based Vision Sensors)》介绍了这种传感器(详见2010年IEEE国际电路与系统研讨会(ISCAS)的会议记录第2426-2429页)。相对于常规的摄像机而言,基于事件的视觉传感器的优点是可消除冗余、减少延迟时间并增加动态范围。

[0017] 这种视觉传感器的输出,就每个像素地址而言,都包括代表产生场景时场景反射系数变化的一组异步事件。

[0018] 传感器的每个像素都是独立的并且检测由最后事件所发出的高于阈值的亮度变化(例如,对比度大于亮度对数的15%)。在亮度变化超出阈值集的情况下,由像素根据亮度增加或者减少(DVS传感器)来产生ON事件或者OFF事件。某些异步传感器将所检测到的事件与光亮度(ATIS传感器)的测量值相互关联。

[0019] 由于传感器没有像常规的摄像机那样以时钟进行取样,所以事件的排序就可以非常高的时间精确度(例如,约为1 $\mu$ s)来考虑。如果将这种传感器用于重建一组图像,则可以达到数千赫的图像率,与常规摄像机的几十赫形成对比。

[0020] 这类摄像机的高时间精确度使之有可能更加充分地利用视频的时间信息。

[0021] 然而,处理来自这些传感器的事件可能比较复杂,因为事件严格遵循时间(t)和空间(x,y)方面的守时概念。于是,处理事件以及分析事件都会比较困难。

[0022] 因此,需要制造简单的装置,可以操作所述装置,以便对来自于异步传感器的信号进行相关分析。

[0023] 分析必须具体包括空间维度和时间维度,从而有助于在这种传感器所传送的极多数量的事件之中识别时空特征,而不会丢失传感器的动态。

[0024] 为此目的,本发明提出一种处理异步信号的方法,所述异步信号是由光传感器产生的,该传感器所具有的像素矩阵对应于场景所相对的位置,所述方法包括:

[0025] -接受来自光传感器的异步信号,对于矩阵的每个像素而言,所述异步信号都包括来自所述像素的连续事件;

[0026] -利用在异步信号所接收到的事件的活动简档(activity profile)来分析异步信号。

[0027] 活动简档,对于传感器的每个像素而言,至少包括活动数值,所述活动数值由所述像素的连续事件中的最近事件之后所经过的时间作为函数而减少。

[0028] 可将像素的“活动简档”视为时间函数的曲线,其数值至少代表由该像素所接收到的最后一个事件的时间(有可能以指定极性对其进行过滤)。

[0029] 所创建的活动简档由此形成分析工具,用于通过局部地保持事件的时空结构来汇总事件的出现。

[0030] 已经观察到活动简档的形态可用于表示传感器所观察的场景中存在某些基本形式。

[0031] 接受这些改变异步事件的简档并因此保持着传感器的动态。

[0032] 一组活动简档形成所谓的传感器“新鲜度卡”。

[0033] 通过利用活动简档分析可以开发很多信号处理应用程序。按照非限制性的方式,可将其用于:

[0034] -形状识别:参照解释预定义基本形状的预定义运动的典型简档库。如果分析检测到活动简档与库的典型简档相似,则进行形状识别,并且对该形状行进速度的表示进行估算;

[0035] -估算场景中的运动:如果第一次检测到活动简档,那么就在相邻区域活动简档(或者非常相似的活动简档)的较短时间里,可以确定传感器视场中物体行进速度的估值;

[0036] -估算图像(立体)中的深度:如果在来自捕捉相同场景的两个传感器的两幅图像中检测到相同(或相似)的活动简档,就有可能认为与活动简档相关的像素与所捕获的场景的相同点相对应,并因此确定该点的深度;

[0037] -以及其它,等等。

[0038] “异步信号”可以是来自指定视网膜传感器的一组事件或者是这些事件的子集(空间子集:仅限于某些像素;和/或时间子集:限于指定时间段)。

[0039] 此外,活动简档可以根据来自所述像素的连续事件中最近事件起始所经过的时间以指数方式减少。

[0040] 该指数下降使之能够更好地检测到由传感器所接收到的最后的事件。

[0041] 在一个特殊实施例中,活动简档可进一步作为来自所述像素的连续事件中最近事件之前的事件起始之后所经过时间的函数。

[0042] 乃至可以把从该像素接收的所有事件都考虑在内,以便确定相关联的活动简档。

[0043] 本发明还提出一种识别特别适合视网膜传感器的形状的方法,并且利用前述异步信号的处理。

[0044] 本发明还可以涉及到识别形状的方法,包括:

[0045] /a/执行处理前述异步信号的方法;

[0046] /b/当接收到事件时,针对所述传感器的每个当前像素:

[0047] /b1/确定所述当前像素的背景,所述背景定义为其位置距离所述当前像素小于预定距离的像素的一组活动简档,

[0048] /b2/从一组预定典型背景中确定典型背景,使从一组典型背景中所确定的典型背景与在步骤/b1/中所确定背景的距离最小,

[0049] /c/根据步骤/b2/中所确定的典型背景的出现次数,来确定识别标志;

[0050] /d/通过所述识别标志与典型识别标志基础的比较,来确定形状。

[0051] 距离可以是数学术语的距离。就此而论,距离可以是欧几里德距离、曼哈顿距离、闵可夫斯基距离、切比雪夫距离或者任何其它距离。

[0052] 通常,可以把“背景”表示为三维空间中的平面(两条轴线表示像素的坐标,一条轴线没有维度(作为时间函数))。

[0053] 背景使之能够在空间和时间上认识所考虑的事件(即当前事件)的附近环境。可将背景视为与像素坐标相关的一组数值,相对于当前像素而言,所述像素坐标的位置与当前事件从中而来的像素(称为“当前像素”)相间隔的距离小于预定距离。

[0054] 通常,可以把一组分量生动地表示为“柱状图”。在其余说明中也采用术语“柱状图”或“识别标志”。

[0055] 在另一个实施例中,可以针对所采用的典型背景将分层模型考虑在内。

[0056] 本发明本身涉及一种识别形状的方法,所述方法包括(确定典型背景的分层模型,每个典型背景都与分层模型的多个等级相关联):

[0057] /a/将传感器异步信号的事件用作当前事件,并且将分层模型的第一级别的典型背景用作当前典型背景;

[0058] /b/执行适用于处理异步信号的方法,如前文针对当前事件所述;

[0059] /c/当接收到当前事件时,对于所述传感器的每个当前像素而言:

- [0060] /c1/确定所述当前像素的背景,对于与所述当前像素相间隔的距离小于预定距离的像素,将所述背景定义为在步骤/b/中所使用的一组活动简档,
- [0061] /c2/从当前典型背景(504)中确定典型背景( $C_k$ ),所确定的典型背景在所有当前典型背景之中具有与步骤/c1/中所确定的背景的最小距离,
- [0062] /c3/生成以在步骤/c2/中所识别的典型背景作为函数的事件;
- [0063] /d/如果尚未使用分层模型的级别:
- [0064] -将在步骤/c3/中所生成的事件用作当前事件,
- [0065] -将尚未使用的分层模型中的级别典型背景用作当前典型背景;
- [0066] -再次执行步骤/b/至/e/;
- [0067] /e/如果已经使用了分层模型中的所有级别:
- [0068] -确定(521)以在步骤/c2/中所最后出现的典型背景所出现的次数作为函数的识别标志;
- [0069] -通过比较所述识别标志与典型识别标志的基础来确定(524)形状。
- [0070] 这些分层的层次等级使之有可能进行一步一步地连续识别:使之有可能识别出非常基本形状的第一个层次等级,使之有可能利用较低级别来识别出具有更复杂形状的较高层次等级。
- [0071] 在一个特殊实施例中,背景的确定分别把具有不同极性的事件考虑在内。
- [0072] 例如,极性可以与像素亮度可变的这一事实相对应(例如,在增加的情况下为+1或者在减少的情况下为-1)。极性还可以与针对紧随其后的较低层次等级识别的典型背景相对应。
- [0073] 将具有不同极性的事件考虑在内分别能够增加所确定的背景的相关性。
- [0074] 在这种情况下,可将背景视为与下列内容相关联的一组数值:
- [0075] -像素的坐标,相对于当前像素而言,所述像素坐标的位置与“当前像素”相间隔的距离小于预定距离;以及,
- [0076] -指定极性。
- [0077] 有利的是,在步骤/e/中所采用的距离是巴氏距离或者标准化距离。
- [0078] 计算机程序,且安装在已有设备上并执行上文所述的全部或一部分方法,其本身是有利的。
- [0079] 就此而论,本发明还涉及到一种计算机程序,在通过处理器执行该程序的情况下,所述计算机程序包括执行上文所述的方法的指令。
- [0080] 该程序可以采用任何编程语言(例如,面向对象的语言或其它语言),并且其形式为可说明的源代码、部分汇编代码或完全汇编代码。
- [0081] 下文所述的图5可以形成这种计算机程序的一般算法的流程图。

## 附图说明

- [0082] 本发明的其它特征和优点可通过阅读以下说明得以进一步呈现。以下说明仅处于阐释性目的,必须结合附图阅读,在附图中:
- [0083] -图1是ATIS型异步光传感器的概要图;
- [0084] -图2简要图示了由面对包括旋转星的场景的异步传感器所产生的事件;

- [0085] -图3a和图3b是关于指定像素的事件的“活动信号”的实施例
- [0086] -图3c是说明根据本发明实施例所指定“新鲜度”卡的3D表面的一个示例；
- [0087] -图3d是说明根据本发明实施例的背景的3D表面的一个示例；
- [0088] -图4显示了在返回玩扑克牌时的事件流中事件的出现；
- [0089] -图5是本发明一个可能的实施例的流程图示例；
- [0090] -图6a至图6c是本发明一个可能的实施例的典型背景的示例；
- [0091] -图7显示了实施根据本发明实施例的装置；
- [0092] -图8a至图8c是针对各种待识别形状的检测形状的性能实例。

## 具体实施方式

[0093] 图1显示了ATIS的原理。

[0094] 形成传感器的矩阵的像素101包括两个光敏元件102a、102b,比如二极管,各自分别与电子检测电路103a、103b相连。

[0095] 在由二极管102a所接收的光亮度按照预定量变化的情况下,传感器102a及其电路103a产生脉冲 $P_0$ 。

[0096] 标记该亮度变化的脉冲 $P_0$ 触发与其它二极管102b相关联的电子电路103b。然后,二极管102b一旦接收到指定量的光(光子的数量),该电路103b便产生第一个脉冲 $P_1$ ,然后产生第二个脉冲 $P_2$ 。

[0097] 脉冲 $P_1$ 与 $P_2$ 之间的时间差 $\delta t$ 与刚刚出现脉冲 $P_0$ 之后由像素101所接收到的光亮度成反比。

[0098] 来自ATIS的异步信息包括针对每个像素(104)相组合的两个脉冲串:第一个脉冲串 $P_0$ 表示光亮度变化到检测阈值以外的瞬间,而第二个脉冲串则由脉冲 $P_1$ 和 $P_2$ 所构成,其时间差 $\delta t$ 表示相对应的光亮度或者灰度等级。

[0099] 来自ATIS的矩阵位置 $p$ 的像素101的事件 $e(p, t)$ 包括两种类型的信息:通过脉冲 $P_0$ 的位置所给定的时间信息,指出事件的瞬间 $t$ ,以及由脉冲 $P_1$ 与 $P_2$ 之间的时间差 $\delta t$ 所给定的灰度等级信息。

[0100] 来自像素的事件的位置可以用如图2所示的三维时空表达来设置。在该图中,各个点 $p$ 标记通过按照方框A中用图表示的不变角速度旋转的星的运动在位置为 $\begin{pmatrix} x \\ y \end{pmatrix}$ 的传感器的像素 $p$ 的水平在瞬间 $t$ 所异步生成的事件 $e(p, t)$ 。其中,大部分点分布在大致呈螺旋形的表面附近。此外,该图还显示了与螺旋面相隔一定距离的一定数量的事件,且在无需与星的有效运动相对应的情况下来测定所述事件。这些事件为采集噪声。

[0101] 然后,事件 $e(p, t)$ 可由以下全部信息来定义:

$$[0102] \quad e(p, t) = \begin{cases} p \in C \subset \mathbb{R}^2 \\ pol \\ I(p, t) \end{cases}$$

[0103] 式中: $C$ 为传感器的空间域, $pol$ 极性代表亮度改变的方向(例如,1代表增加或者-1代表减少)以及 $I(p, t)$ 为点 $p$ 的光亮度信号。

[0104] 光亮度信号本身可以是一组组合的脉冲串104,如图1所示。如果 $t_0$ 代表出现事件



的覆盖以及 $t_{e+.u}$ 与 $t_{e-.u}$ 之差表示与亮度变化成反比的数值,则可以利用三个狄拉克 $\delta$ 对亮度进行编码,使得 $I(p, t) = \delta(t - t_u) + \delta(t - t_{e+.u}) + \delta(t - t_{e-.u})$ 。使得位于坐标 $p$ 的像素的亮度信号能够给亮度信息进行暂时编码。

[0105] 可以注意到,传感器的第 $i$ 个事件作为 $ev(i)$ ,则 $v(i)$ 可以通过以下所有信息来确定:

$$[0106] \quad ev(i) = \begin{cases} p_i \in C \subset \mathbb{R}^2 \\ pol_i \\ t_i \\ I_i(p_i, t_i) \end{cases}$$

[0107] 式中: $C$ 为传感器的空间域, $p_i$ 为与第 $i$ 个事件有关的传感器的点, $pol_i$ 用极性来表示事件的类型(例如,第 $i$ 个事件的亮度变化的方向,例如,1表示增加或者-1表示减少), $t_i$ 为出现第 $i$ 个事件的事件,以及 $I_i(p, t)$ 为点 $p_i$ 的光亮度信号(如果该数值可供使用)。

[0108] 为了以更简单的方式来控制事件,可以针对每个像素 $p$ 以及针对每个极性 $pol$ 来定义表示事件的“新鲜度”的函数 $S$ , $S(p, pol)$ 至少为针对该像素并且具有该极性 $pol$ 的最后事件所出现的时间 $t$ 的函数。

[0109] 可将函数 $S$ 定义为针对于指定像素 $p$ 以及指定极性 $pol$ 的每个事件

$$ev_p(i) = \begin{cases} p_i = p \\ pol_i = pol \\ t_p \end{cases} \text{ 而言,原函数 } S_{prim}(p, pol, t) = \begin{cases} \max(h + \theta \cdot (t_p - t), 0) & \text{si } t \geq t_p \\ 0 & \text{sinon} \end{cases}$$

在指定瞬间 $t$ 的和,式中: $h$ 是预先确定的数值, $\theta$ 是与原函数的递减速度相对应的预先确定的系数。

[0110] 原函数的“和”在数学上可看作以下内容的卷积:

$$[0111] \quad \text{—原函数 } S_{prim}(p, t) = \begin{cases} \max(h - \theta \cdot t, 0) & \text{si } t \geq 0 \\ 0 & \text{sinon} \end{cases} \text{ (或者更普通而言,任何递减函数),}$$

$$[0112] \quad \text{—信号包括发生事件 } ev_p(i) = \begin{cases} p_i = p \\ pol_i = pol \\ t_p \end{cases} \text{ 的各个时间 } t_p \text{ 的狄拉克。}$$

[0113] 出于阐释性目的,图3a显示关于传感器的三个像素 $p_1$ 、 $p_2$ 和 $p_3$ (以及关于指定极性值 $pol$ )的三种可能性活动信号 $t \rightarrow S$ 。

[0114] 在不存在事件的情况下, $S(p_1, t)$ 、 $S(p_2, t)$ 或 $S(p_3, t)$ 的数值都为零。然而,在出现关于像素 $p_1$ 的极性事件 $pol$ (例如,310)的过程中, $S(p_1, t)$ 取预定阈值(在此为 $h$ ,该数值 $h$ 能够单一的)。

[0115] 然后,在该事件朝0移动之后,活动信号 $S(p_1, t)$ 的值逐渐减少。

[0116] 这一点也适用于像素 $p_1$ 的事件311、像素 $p_2$ 的事件312或者像素 $p_3$ 的事件313/314,

[0117] 如果活动信号 $S$ 的减少在此是线性的,则可以提供任何类型的减少,比如指数下降:

$$[0118] \quad h \cdot e^{-\frac{t-t_p}{\tau}}$$

[0119] 该指数下降可以通过图3b来阐释(见曲线3b和事件320)。

[0120] 此外,在出现针对所考虑的像素的事件的过程中(例如在此为 $p_4$ ),相对于 $h$ 的数值而言,函数 $S$ 的数值不可忽略(例如,事件321暂时接近事件322)。

[0121] 在一个实施例中,在出现随后的事件322的过程中,可将活动信号 $S$ 的数值设为紧接着在事件322之前的 $S$ 的当前值(即 $h_0$ )和 $h$ 之和(有可能是加权和)。就此而言,曲线 $S$ 的下降可以从图3b所示的 $h+h_0$ 值开始。此外,可以使 $h+h_0$ 的数值的上限设置为预定值 $h_1$ (即最小值( $h_1, h+h_0$ ))。

[0122] 在另一个实施例中,在出现随后的事件322的过程中,将曲线 $S$ 的数值设置为值 $h$ ,无论 $h_0$ 的数值是多少(即忽略最后事件之前的事件(即随后的事件))。在这另一个实施例中,被称为定义如下的“最后的事件的时间”的时间可以定义为:

[0123]  $T(p, pol, i) = \max(t_i) \mid j < i$

[0124] 或者

[0125]  $T(p, pol, t) = \max(t_i) \mid t_i < t$

[0126] 式中: $t_i$ 为针对具有极性 $pol$ 的像素 $p$ 发生的事件的时间。

[0127] 在概念上讲, $p \rightarrow T(p, pol, t)$ 定义了刚好在参考时间(即 $t$ )之前暂时发生的相同极性的最后事件的时间的卡片。

[0128] 然后,在这另一个实施例中可以将 $p \rightarrow S(p, pol, t)$ 定义为该时间集合 $T(p, pol, t)$ 的函数。

[0129] 例如, $p \rightarrow S(p, pol, t)$ :

[0130] 
$$p \rightarrow S(p, pol, t) : \begin{cases} h \cdot e^{-\frac{t-T(p, pol, t)}{\tau}} & \text{si } t \geq T(p, pol, t) \\ 0 & \text{sinon} \end{cases}$$

[0131] 式中: $\tau$ 和 $h$ 是预定时间常数( $S$ 可以是时间 $t$ 在包括下限 $T(p, pol, t)$ 区间内的任意递减函数)。

[0132] 代表像素事件“新鲜度”的像素的卡片 $S$ 的形成是有益的,因为它使之能够连续并且简单地表示不连续的概念(即事件)。如此创建的卡片使之能够令事件的表达在简单的理解域中进行变换。

[0133] 然后,其形成也可简化事件的操作和比较。

[0134] 一旦确定了像素卡片 $p \rightarrow S(p, pol, t)$ ,则可以根据坐标 $p$ 创建关于时间 $t$ 以及关于极性 $pol$ 的固定值的 $S$ 的幅度的3D图表(见图3c)。

[0135] 当然, $S$ 可以包括 $N$ 组独立值(即 $((p, pol) \rightarrow S(p, pol, t))$ , (如果有 $N$ 个可能性的极性数值)每个可能性极性数值 $pol$ 都有一组独立值。

[0136]  $S_p$ 被称为像素 $p$ 的“背景”,关于像素 $q$ 的数值 $(q, pol) \rightarrow S(q, pol, t)$ 的集合在像素 $p$ 的附近(即,位于像素 $p$ 的预定距离,例如,数学术语的距离,例如,呈以像素 $p$ 为中心、边长为 $2R+1$ 的正方形)。为了形象地表示背景(包括多个可能性极性值),可以并置关于多个可能值 $pol$ 的 $q \rightarrow S(q, pol, t)$ 的多个表达。

[0137] 例如,图3d可以表示像素 $p$ 的“背景”350:平高线351代表函数 $q \rightarrow S(q, 1, t)$ 的数值,平高线352代表函数 $q \rightarrow S(q, -1, t)$ 的数值(当然,平高线可呈任何形状,平高线的选择仅仅使之能够有助于查看)。同样,点353和354对应于相同像素且不同极性的函数的数值。

[0138] 可以定义针对任何像素 $p$ 的背景 $S_p$ 。

[0139] 为了能够表示各种可能性背景的特征,就有可能定义被称之为“典型”的背景。

[0140] 这些典型的背景都可以预定或者根据由D.Ballard和J.Jehee编写的《皮层反馈电路中符号量的动态编码 (Dynamic coding of signed quantities in cortical feedback circuits)》(详见2012年《心理学前沿 (Frontiers in Psychology)》第3卷编号254)所提供的算法或者利用其它方法(例如“k-means”方法)来确定。

[0141] 关于针对像素p所识别的各个背景 $S_p$ ,可将典型的背景与像素p相关联,其中所述典型背景与最接近 $S_p$ 的典型背景相对应。背景 $S_p$ 与典型背景之间的距离,例如,可以通过计算在相同像素p以及相同极性pol的像素的背景 $S_p$ 的数值与典型背景的数值之间的欧几里德距离之和来确定。该距离还可以是平方欧几里德距离之和的函数。

[0142] 如果所计算的距离高于一定的预定阈值,则可以不将任何典型背景与像素p相关联。

[0143] 将这些典型背景 $\{C_k\}$ 定义在有限空间域内,如上文针对背景 $S_p$ 所述(例如,在边长为 $2R+1$ 的正方形上)。

[0144] 还可以定义典型背景的多个层次等级,各个层次等级m定义多个典型背景 $\{C_{k_m}\}$ 。这种层次等级的益处可参照图5阐述作详细的说明。

[0145] 出于阐释性目的,图6a显示了典型背景的第一个层次等级(即,关于极性为1(每个典型背景的左侧区域)或者关于极性-1(每个典型背景的右侧区域)的尺寸 $5 \times 5$ 的像素)。

[0146] 此外,在指定瞬间t且在与每个像素p相关联的典型背景 $\{C_k\}$ (或者更普遍而言关于固定层次等级m的 $\{C_{k_m}\}$ )的识别之后,就有可能计算出针对所有可能性像素p的各个典型背景的出现次数。

[0147] 所计算出的出现次数使之有可能形成表示事件流特征的识别标志/柱状图 $\mathcal{H}$ (纵坐标为出现的次数,横坐标为所识别的典型背景的指数)。

[0148] 表征形状识别事件流的特征还可以利用其它方法,比如回声状态网络或循环神经网络。

[0149] 出于阐释性目的,图4显示了在返回玩包括草花、方片、红桃和黑桃的扑克牌时的“ON”事件(白色点)以及“OFF”事件(黑色点)的出现。还可以在这些事件流下面展示分别与各个事件流(a)、(b)、(c)和(d)相关的柱状图:这些柱状图形成有待于在事件流中识别的所谓的形状“识别标志”。

[0150] 这些柱状图有可能在确定数量的训练过程中构建出(换言之,通过呈现事件流生成柱状图,所述事件呈现相同的形状):就此而言,有可能通过求训练过程中所得到的关于相同形状和/或相同运动的柱状图的平均值来确定“典型柱状图”。

[0151] 一旦确定了这些典型柱状图,则可以通过时间流来再次确定当前柱状图,并且将该柱状图与在训练阶段所确定的典型柱状图进行比较。

[0152] 与当前柱状图距离最近的典型柱状图使之能够识别出与事件流相对应的形状。

[0153] 还有可能返回k最近的柱状图类型(有可能具有与其近距相对应的分数)。

[0154] 可将柱状图 $\mathcal{H}_1$ 与 $\mathcal{H}_2$ 之间的距离计算为两个矢量之间的数学距离,所述矢量针对坐标具有各个典型背景的出现次数:

$$[0155] \quad d(\mathcal{H}_1, \mathcal{H}_2) = \|\mathcal{H}_1 - \mathcal{H}_2\|$$

[0156] 还可以计算标准化距离如下:

$$[0157] \quad d^N(\mathcal{H}_1; \mathcal{H}_2) = \left\| \frac{\mathcal{H}_1}{\text{card}(\mathcal{H}_1)} - \frac{\mathcal{H}_2}{\text{card}(\mathcal{H}_2)} \right\|$$

[0158] 式中： $\text{card}(\mathcal{H}_j)$ 为柱状图 $\mathcal{H}_j$ 的典型背景(即竖线)的数量。

[0159] 巴氏距离也可以用于代替常规的距离：

$$[0160] \quad d^B(\mathcal{H}_1; \mathcal{H}_2) = -\ln \sum_i \left( \left\| \frac{\mathcal{H}_1(i)}{\text{card}(\mathcal{H}_1)} \cdot \frac{\mathcal{H}_2(i)}{\text{card}(\mathcal{H}_2)} \right\| \right)$$

[0161] 式中： $\mathcal{H}_j(i)$ 为柱状图 $\mathcal{H}_j$ 的第*i*个典型背景出现的次数。

[0162] 也可以采用任何其它数学距离。

[0163] 然后,可以认为在事件流中产生了与相关联的典型柱状图相对应的形状。

[0164] 得益于该典型柱状图,所以形状识别本身是有可能的。

[0165] 图5是本发明的可能实施例的流程图的实例。

[0166] 一旦接收到来自异步传感器的并且与时间 $t_i$ 相关联的事件流500的事件*i*,就可以确定或者更新(步骤501)关于传感器每个像素

以及关于pol的各个数值 $S(p, \text{pol}, t_i)$ ,如上文所述:这个步骤使之能够形成或者更新传感器的“新鲜度”卡。

[0167] 关于步骤501,减少S原函数所用的时间常数表示为 $\tau_1$ 。因此,例如,我们可以得到:

$$[0168] \quad S(p, \text{pol}, t_i) = h \cdot e^{-\frac{t_i - T(p, \text{pol}, t_i)}{\tau_1}}$$

[0169] 一旦进行如此确定,则可以针对传感器的每个像素

从在步骤501中预先计算的新鲜度卡中提取出背景 $S_p$ (步骤502):如此提取使之能够将像素q的某些值 $S(q, \text{pol}, t_i)$ 与所考虑的针对指定极性值pol的像素p的空间近距 $N_p$ 相隔离。出于阐释性目的, $N_{p-1}$ 可以界定中心在所考虑的空间位置p周围的正方形或三角形。可以确定空间近距 $N_{p-1}$ ,使得就此提取的背景的尺寸等于第一个层次等级503( $C_{k-1}$ )的典型背景的尺寸,以便与后者作比较。

[0170] 出于阐释性目的,图6a显示了可能性典型背景的第一个层次等级(即,关于极性为1(每一个典型背景的左侧区域)或者极性为-1(每一个典型背景的右侧区域)的尺寸5x5的像素)。

[0171] 比较所确定的背景 $S_{p-1}$ 并且将像素p与第一个层次等级的可能性典型背景 $\{C_{k-1}\}$ 相关联,使之能够识别出最接近 $S_{p-1}$ 的典型背景(步骤504),如上述。

[0172] 在第一个层次等级的可能性典型背景 $\{C_{k-1}\}$ 中识别出最接近的典型背景 $C_{k-1\text{prox}}$ ,使之能够针对像素p生成表示与该像素p相关联的典型背景的事件 $ev_1$ (步骤505):

$$[0173] \quad ev_1(p) = \begin{cases} t \\ C_{k-1\text{prox}} \end{cases}$$

[0174] 式中: $t$ 为当前时间。

[0175] 如果没有识别出最近的典型背景(见上文),则不生成事件。

[0176] 生成这些事件 $ev_1(p)$ 还形成异步事件流(506)。

[0177] 然后,可以按照与来自传感器的事件500相似的方式来处理这些事件506。

[0178] 同样,一旦接收到事件流506中与时间 $t_i$ 相关联的每个事件*j*,便可以确定或更新(步骤511)新的“新鲜度”卡的数值(即,关于第二个层次等级的菜单),使之针对传感器每个



像素 $p$ 以及针对 $pol$ 的每个数值都具有数值 $(p, pol) \rightarrow S_2(p, pol, t_j)$ , (已知关于该步骤511 $pol$ 的数值代表在针对第一个层次等级的步骤504过程中识别的典型背景, 即图6a的框架中可能存在的4个典型背景, 每个典型背景具有两个极性), 如上所述。

[0179] 关于步骤511, 减少原函数所用的时间常数表示为 $\tau_2$ , 使 $\tau_2 \geq \tau_1$ 。于是, 例如, 我们可以得到:

$$[0180] \quad S_2(p, t_j) = h.e^{-\frac{t_j - T(p, t_j)}{\tau_2}}$$

[0181] 一旦执行了如此确定, 则可以针对传感器的每个像素 $p$ 从上文所述的在步骤511中所计算的新鲜度卡中提取出背景 $S_{p,2}$  (步骤512): 如此提取使之能够将关于具有空间近距 $N_{p,2}$ 的像素 $q$ 的某些值 $S_2(q, pol, t_i)$ 与所考虑的针对指定极性值 $pol$ 的像素 $p$ 相隔离。出于阐释性目的,  $N_{p,2}$ 可以界定中心在所考虑的空间位置 $p$ 周围的正方形或者矩形。可以界定空间近距 $N_{p,2}$ , 使之所提取出的背景的尺寸等于第二个层次等级513 ( $C_{k,2}$ ) 的典型背景的尺寸, 以便与后者进行比较。

[0182] 出于阐释性目的, 图6b显示了八种可能性典型背景的第二个层次等级。

[0183] 如果第一个层次等级的每个典型背景都有两个区域 (极性-1 (OFF) 的事件的一个区域以及极性1 (ON) 的事件的一个区域, 见图6a), 则第二个层次等级的典型背景具有与紧接着下一个层次等级 (即, 在此为第一个层次等级, 见图6b) 的典型背景一样多的区域 (即, 4个)。

[0184] 此外, 第二个层次等级的典型背景可以使得空间近距 $N_{p,2}$ 可以界定一个区域, 该区域大于空间近距 $N_{p,1}$ 所界定的区域 (例如, 如果由边长为 $2R_1+1$ 的正方形来界定 $N_{p,1}$ , 则可以通过边长为 $2R_2+1$ 的正方形来界定 $N_{p,2}$ , 使 $R_2 \geq R_1$ )。

[0185] 比较所确定的背景 $S_{p,2}$ 并且将像素 $p$ 与第一个层次等级的可能性典型背景 $\{C_{k,2}\}$ 相关联, 使之能够识别出最接近 $S_{p,2}$ 的典型背景 (步骤514), 如上所述。

[0186] 由此, 在第二个层次等级的可能性典型背景 $\{C_{k,2}\}$ 中识别出最接近的典型背景 $C_{k,2prox}$ , 使之能够针对像素 $p$ 生成表示与该像素 $p$ 相关联的典型背景的事件 $ev_2$  (步骤515):

$$[0187] \quad ev_2(p) = \begin{cases} t \\ C_{k,2prox} \end{cases}$$

[0188] 式中:  $t$  为当前时间。

[0189] 如果未识别出最接近的典型背景 (见上文), 则不生成事件。

[0190] 生成这些事件 $ev_2(p)$ 还形成异步事件流 (516)。

[0191] 然后, 可以按照与事件506相似的方式来处理这些事件516: 如果存在较高层次等级 (测试517, 输出OK, 例如, 用图6c (6 (c) (1) 和6 (c) (2)) 来表示三级的层次等级), 则可以反复进行步骤511至515, 把通过紧接着下一个层次等级所生成的事件看作输入, 并且利用这较高的层次等级的典型背景 (见513)。

[0192] 如果没有较高的层次等级, 则可以计算针对所有像素 $p$ 以及针对固定时间 $t$ 所识别出的典型背景 (513) 出现的次数。如上文所述, 如此计算使之能够确定代表所识别出的典型背景出现次数的柱状图 $\mathcal{H}$  (步骤521)。

[0193] 利用在步骤513中所确定的柱状图, 可以计算典型柱状图 (523) 之间的距离, 所述

柱状图代表形状和/或运动并由此确定最近的典型柱状图(步骤522)。

[0194] 由于这种确定的典型柱状图,有可能确定与所述柱状图相关联的形状和/或运动(步骤524)并因此返回该形状和/或该运动(525)。

[0195] 当然,如果图5的流程图显示了在分析至少一个第二个层次等级的典型背景之后步骤520至524的执行,便可以在步骤501至505之后完整地执行步骤520至524(即,采用单个层次等级的分析)。

[0196] 图7显示了执行根据本发明实施例的装置。

[0197] 在该实施例中,装置包括计算机700,所述计算机包括存储器705,用于储存使之能够执行所述方法的指令、关于所接收的事件流的数据以及用于执行上文所述方法的各个步骤的暂存数据。

[0198] 计算机进一步包括电路704。例如,该电路可以是:

[0199] -能够以计算机程序的形式编译指令的处理器,或者,

[0200] -电子卡,在其中以硅描述了本发明方法的步骤,或者,

[0201] -可编程电子芯片,比如FPGA芯片(场可编程门阵列)。

[0202] 计算机包括输入接口703,用于从传感器接收事件,以及输出接口706,用于提供在事件流中所识别的形状707。最后,为了方便与用户相互作用,计算机可包括屏幕701和键盘702。当然,键盘是可选的,尤其是在形式为触摸屏平板电脑的计算机的框架中更是如此。

[0203] 图8a是形状(草花、方片、红桃和黑桃)检测的性能示例。

[0204] 各行(见纵坐标801)都与关于已经知悉的具体形状的数据相对应(即,典型柱状图,见上文)。

[0205] 各列(见横坐标802)都与关于已经识别的具体形状的数据相对应(即,待识别的柱状图)。

[0206] 行与列的交叉(例如,与集合803相对应)使之能够查看典型柱状图与针对形状的多次呈现的待识别的柱状图之间的距离(在此,9次呈现形状,由垂直虚线将这些呈现相隔开):

[0207] -白条(左侧条)与常规距离相对应;

[0208] -灰条(中间条)与标准化距离相对应;

[0209] -黑条(右侧条)与巴氏距离相对应。

[0210] 就此所示的三种距离(分别为常规距离、标准化距离和巴氏距离)分别为这些形状的性能的94%、100%以及97%的性能。

[0211] 图8b是形状(数字和大写字母)检测的性能示例。

[0212] 各行(见纵坐标804)都与关于已经知悉的具体形状的数据相对应(即,典型柱状图,见上文)。

[0213] 各列(见横坐标805)都与关于待识别的具体形状的数据相对应(即,待识别的柱状图)。

[0214] 行与列的交叉(例如,与集合806相对应)使之能够查看典型柱状图与针对形状的呈现的待识别的柱状图之间的距离:

[0215] -白条(左侧条)与常规距离相对应;

[0216] -灰条(中间条)与标准化距离相对应;

- [0217] -黑条(右侧条)与巴氏距离相对应。
- [0218] 就此所示的三种距离(分别为常规距离、标准化距离和巴氏(Bhattacharyya)距离)全部为这些形状的100%的性能。
- [0219] 图8c是形状(面部)检测的性能示例。
- [0220] 各个表格(807、808和809)都呈现利用已知人脸(纵坐标)针对该人脸19次呈现所识别出的所显示的人脸(横坐标)的出现次数。
- [0221] 表格807利用常规距离进行人脸识别(识别率:37%)。
- [0222] 表格808利用标准化距离进行人脸识别(识别率:78%)。
- [0223] 表格808利用巴氏距离进行人脸识别(识别率:79%)。
- [0224] 此外,图5所示的功能图是程序的典型示例,可以通过上文所述的装置执行所述程序的某些指令。就此而言,图5可以与根据本发明的计算机程序的一般算法的流程图相对应。
- [0225] 当然,本发明不仅限于上文作为示例所述的实施例;本发明还延伸到其它可选方案。
- [0226] 还可能其它实施例。
- [0227] 例如,图6a至图6c的典型背景可以是任何维度的。
- [0228] 而且,说明主要提及来自光/视频传感器的事件来源。然而,所述的本发明还可推广到任何组的事件来源,比如,在这种异步模式下起作用的传感器的压力网络,所述传感器的压力网络的空间布置可与视频传感器相似。

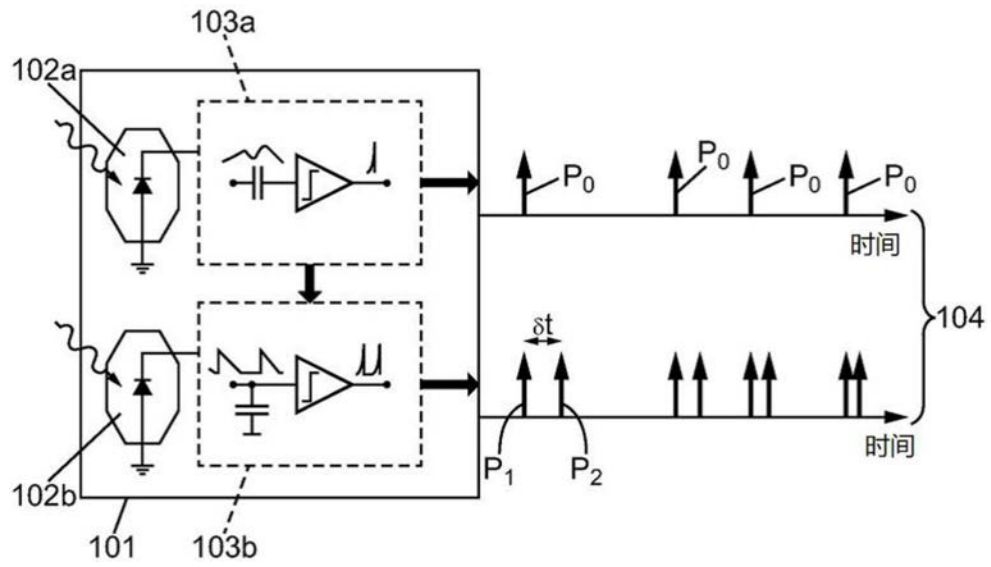


图1

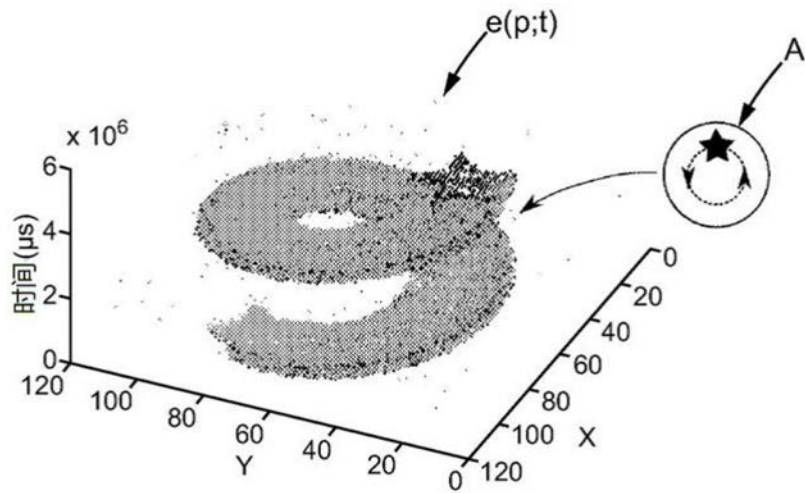


图2



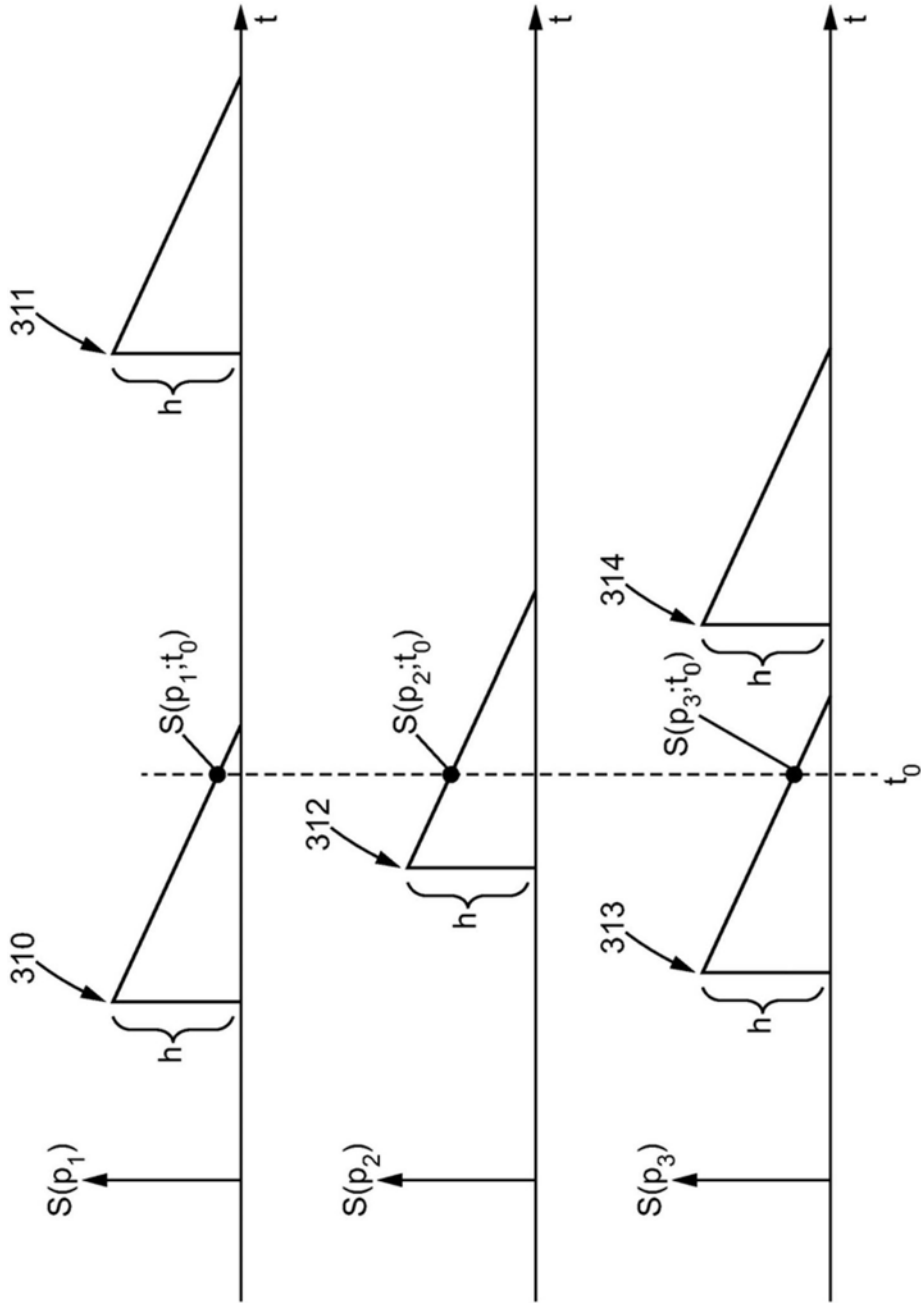


图3a

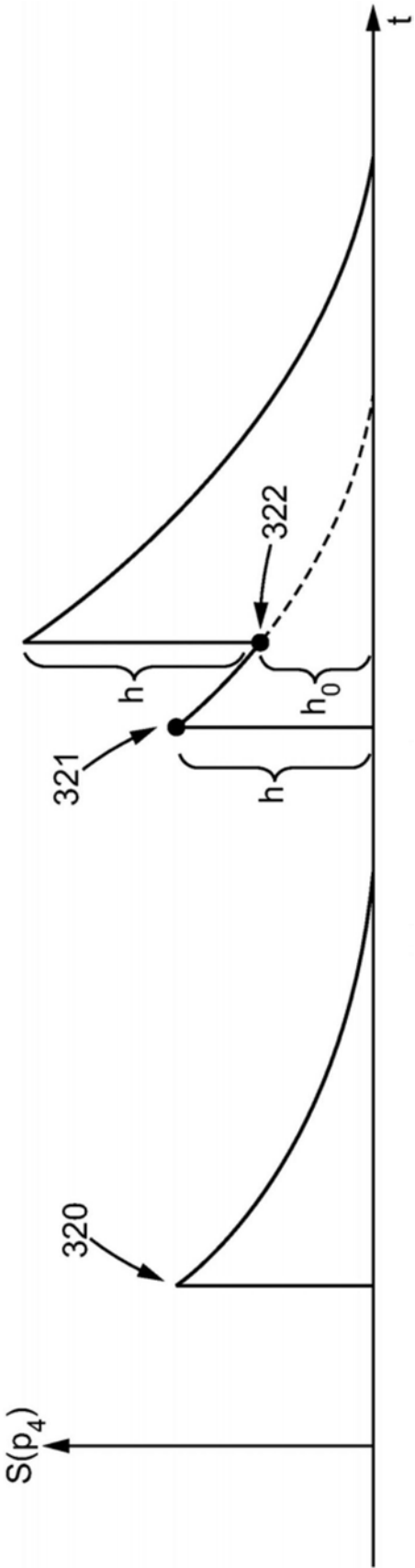


图3b

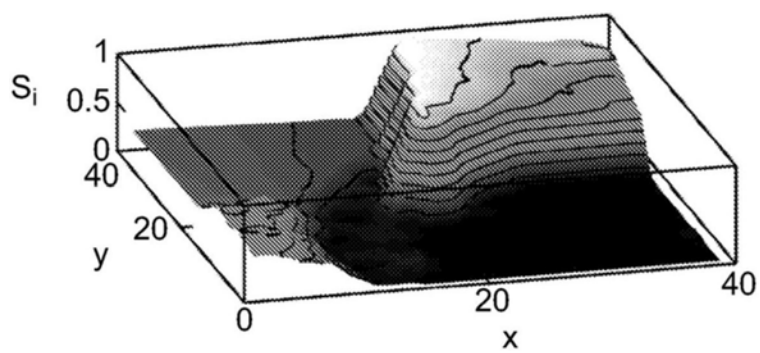


图3c

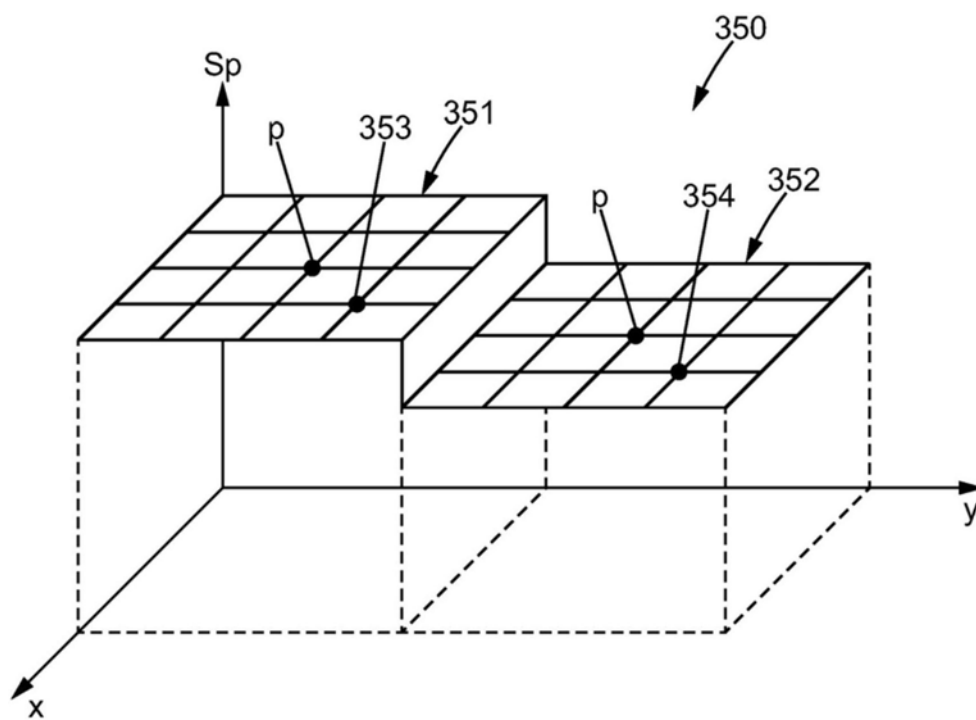


图3d

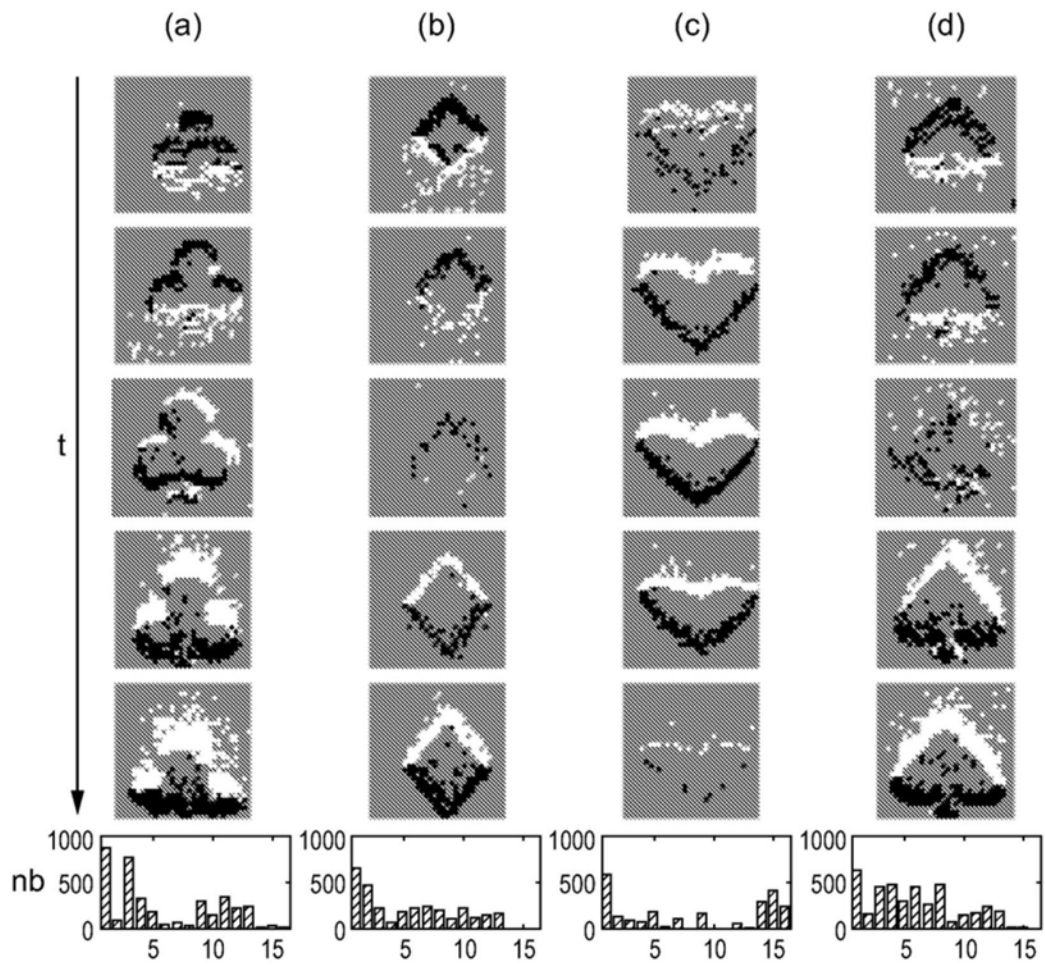


图4

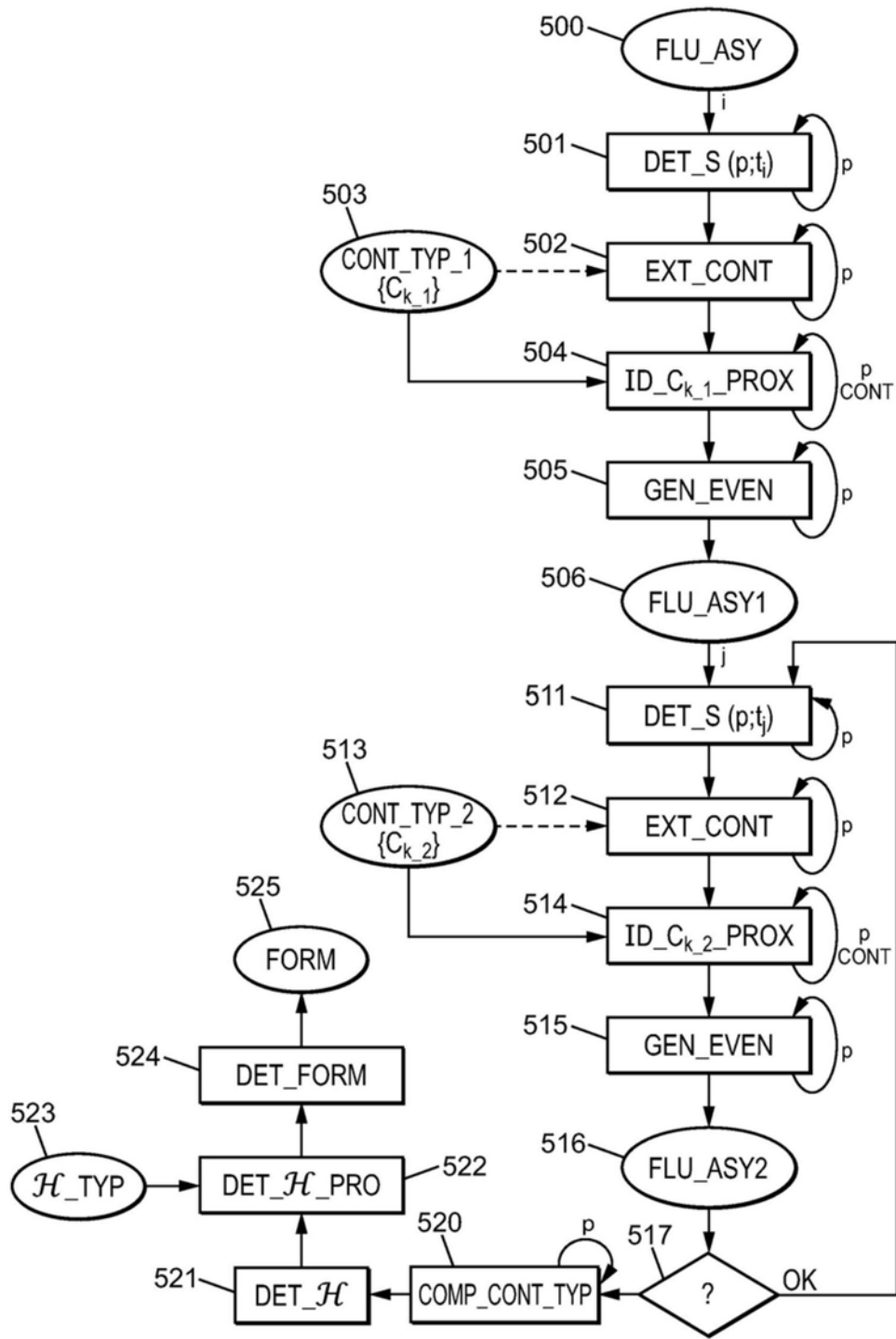


图5

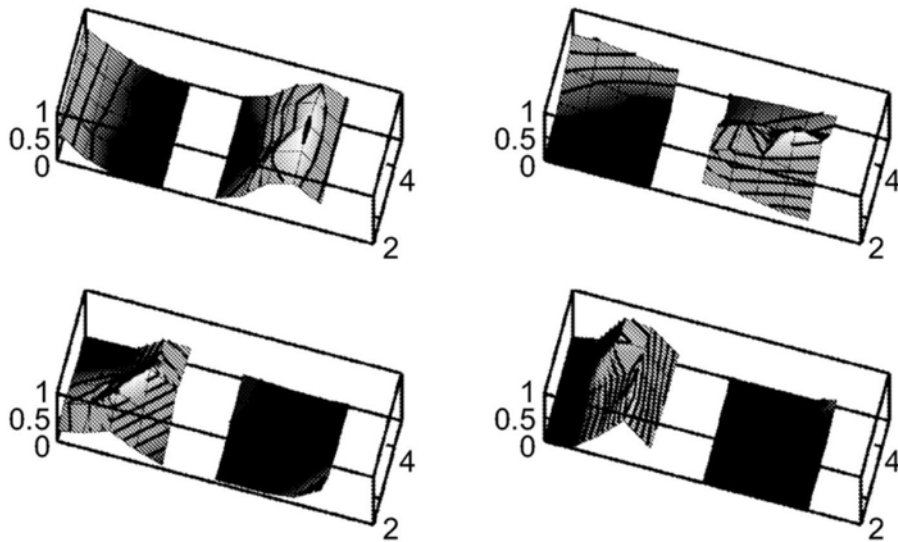


图6a

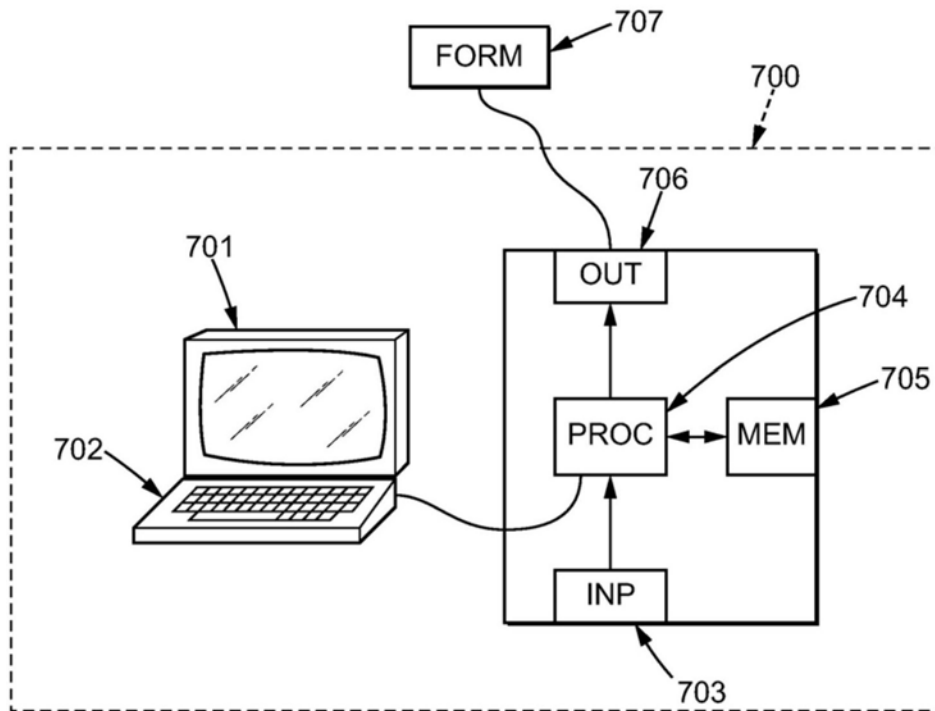


图7

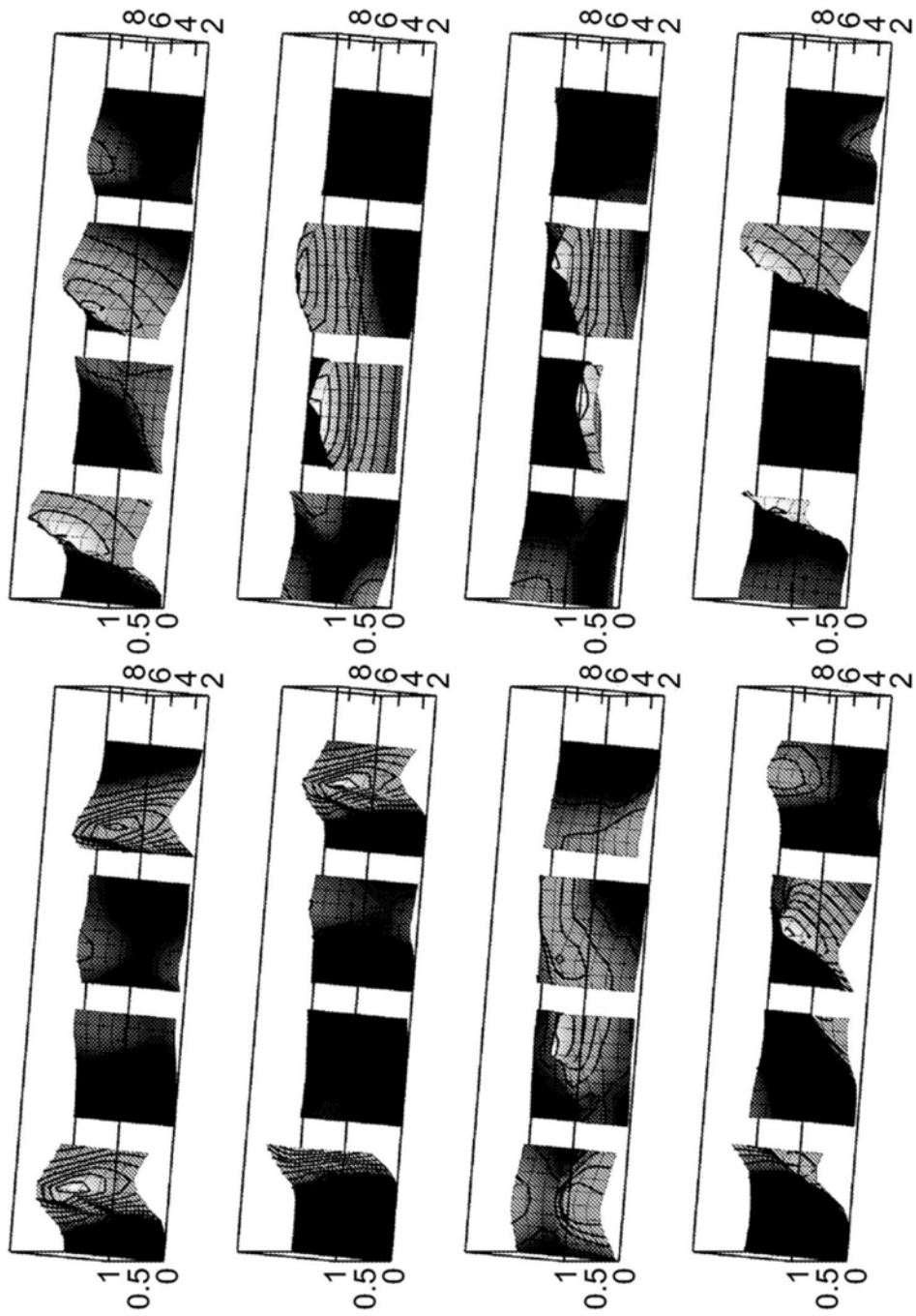


图6b

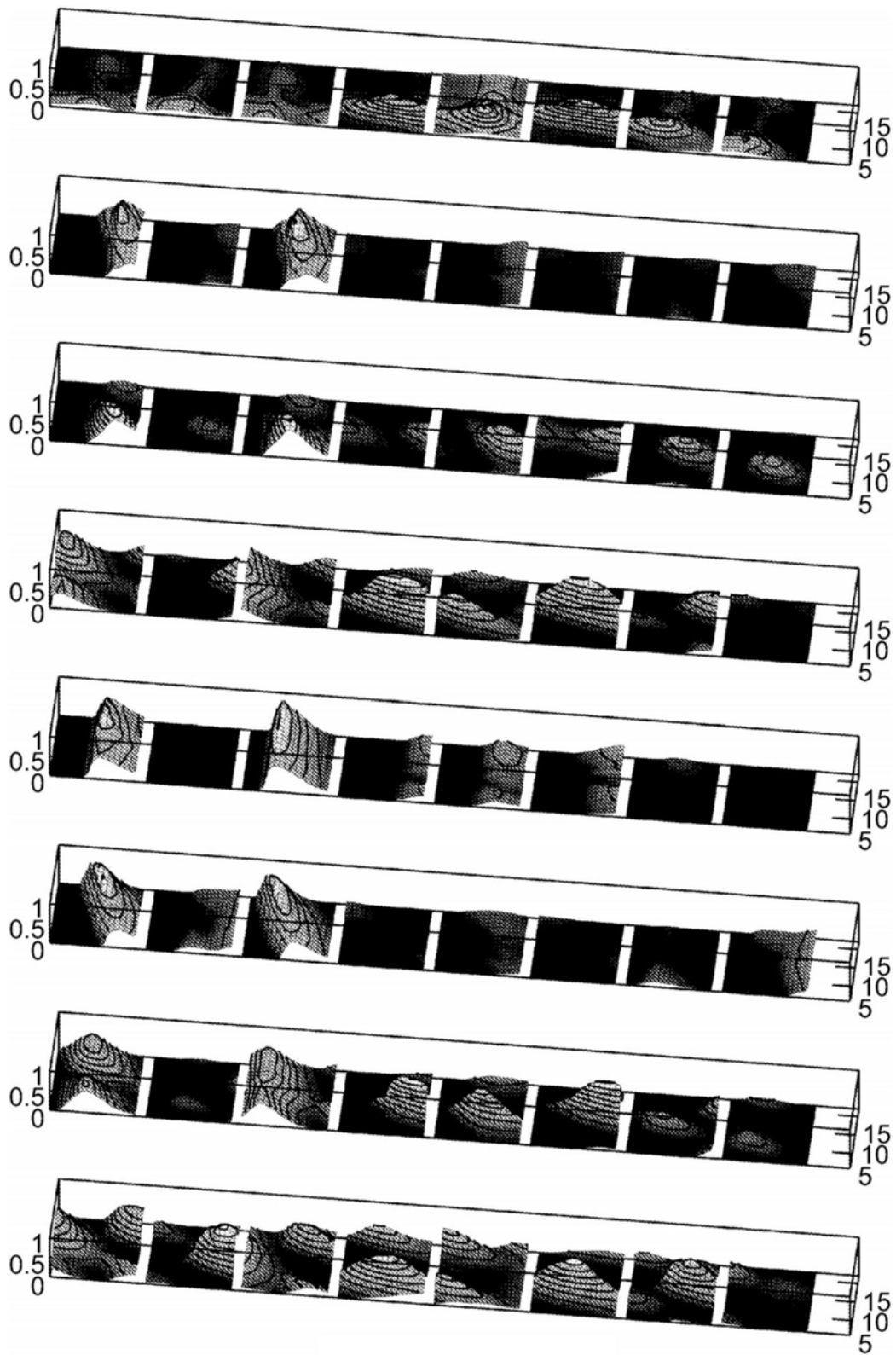


图6c (1)



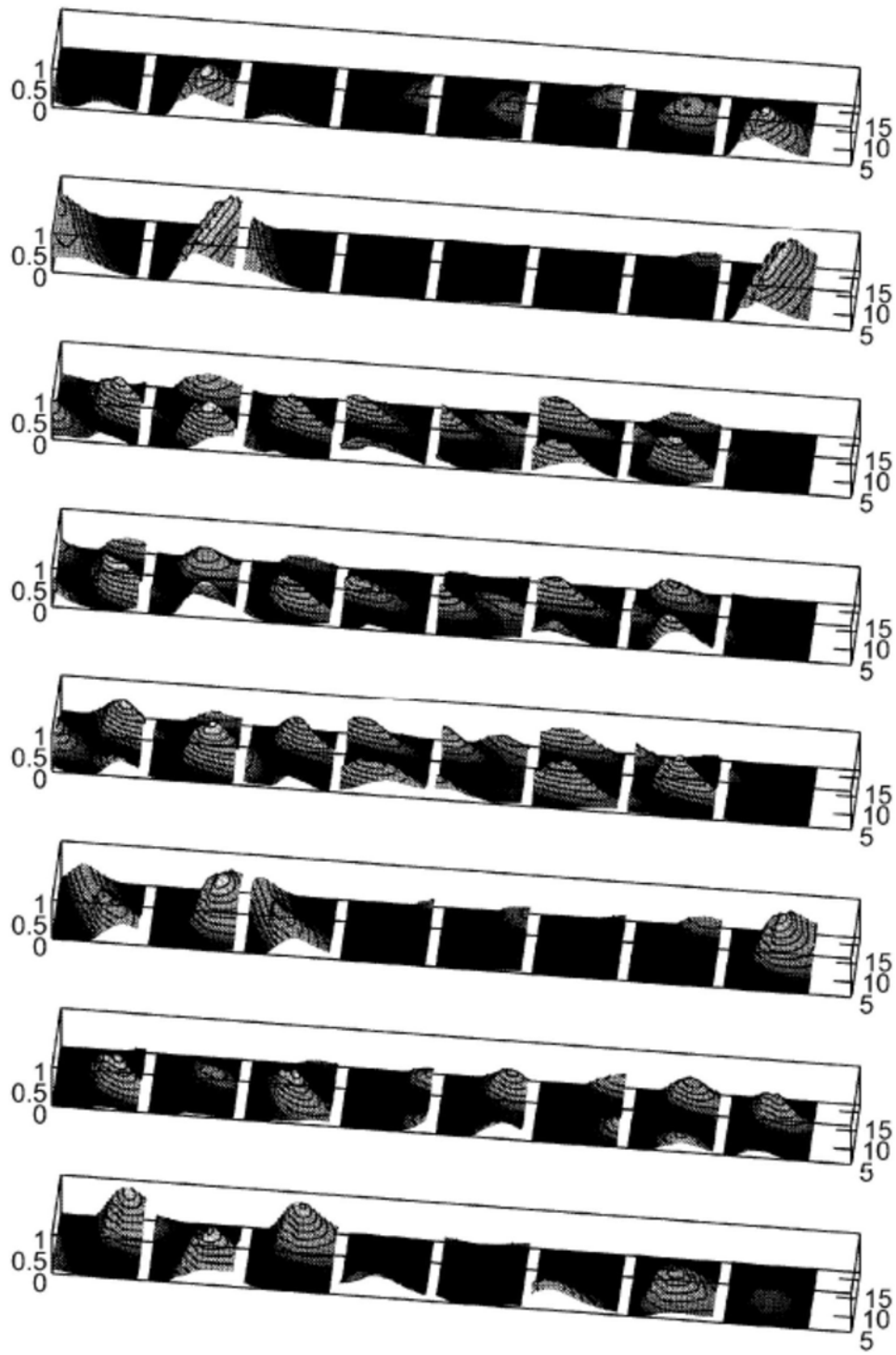


图6c (2)

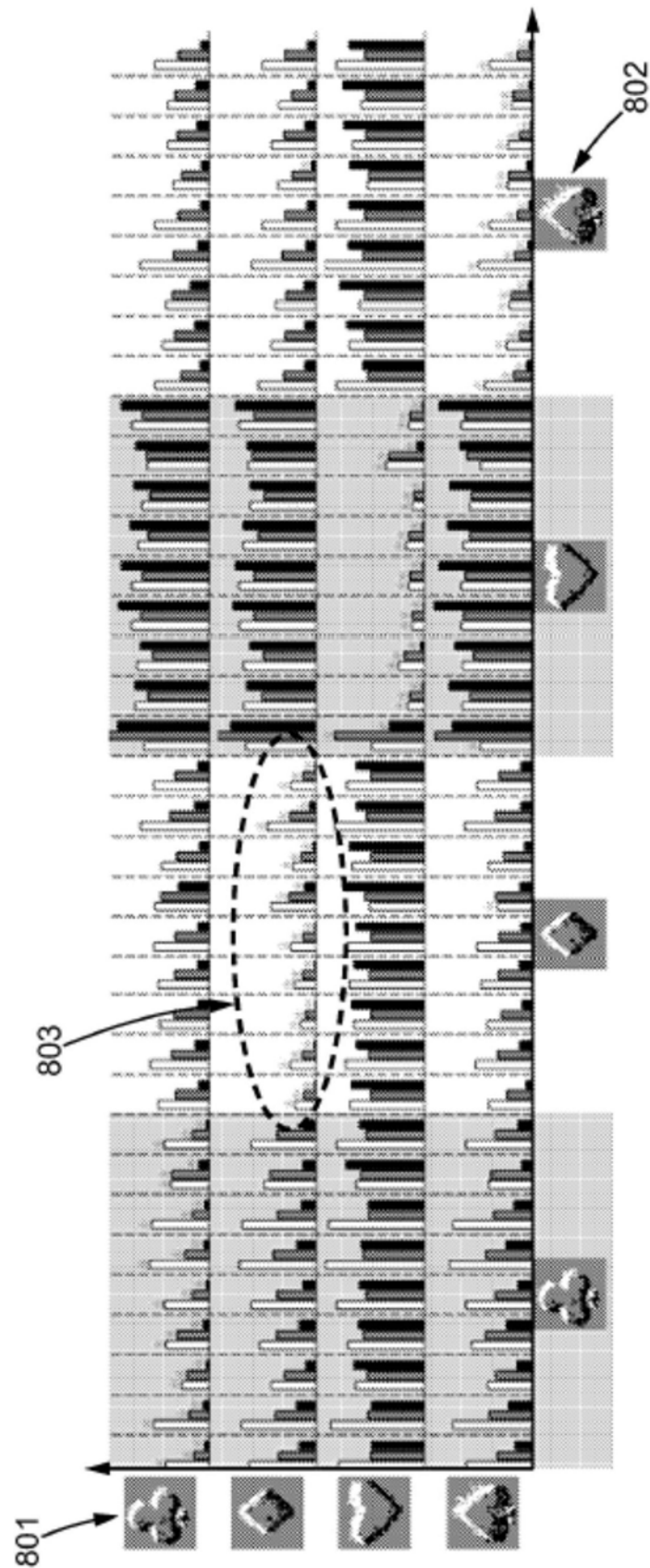


图8a

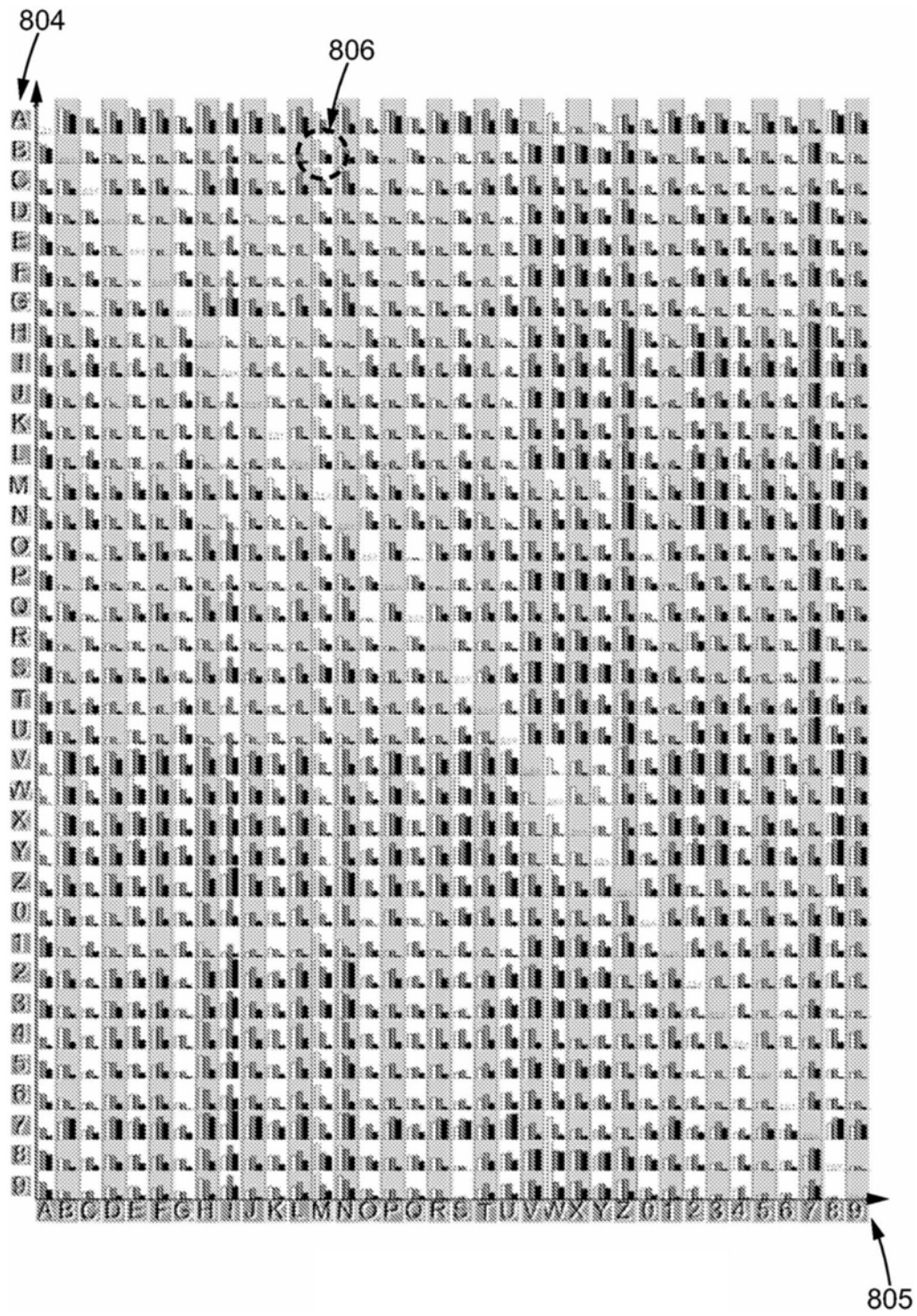


图8b

807 →

(a)							
	11		1	16			3
		5					3
		1					1
	6	6	9	3	1		5
	2	5	9		10		6
						19	
		2			8		1

808 ←

(b)							
	14	2	4				
		15	2	1	1		1
			9	2			
	3			14			
			1		16		
	2	2	1	2	2	19	1
			2				17

809 →

(c)							
	14						
		17			2		2
		1	13	4			
	3		4	13	1		
					16		4
	2	1	2	2		19	
							13

图8c