HAL
open science

Modélisation de l'ambiguïté d'une multi-stabilité visuelle

Kevin Parisot, Alan Chauvin, Ronald Phlypo, Steeve Zozor

To cite this version:

Kevin Parisot, Alan Chauvin, Ronald Phlypo, Steeve Zozor. Modélisation de l'ambiguïté d'une multi-stabilité visuelle. GRETSI, Aug 2019, Lille, France. hal-02325594v1

HAL Id: hal-02325594

https://hal.science/hal-02325594v1

Submitted on 22 Oct 2019 (v1), last revised 8 Nov 2019 (v2)

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.

Modélisation de l'ambiguïté d'une multi-stablilité visuelle

Kevin PARISOT[1,2], Alan CHAUVIN[2], Ronald PHLYPO[1], Steeve ZOZOR[1]

[1]GIPSA-lab, Grenoble Image Parole Signaux Automatic
11 rue des Mathématiques, BP46, F - 38402 Saint Martin D'Hères Cedex, France

[2]LPNC - Laboratoire de Psychologie et NeuroCognition
BSHM - 1251 Av Centrale, CS40700, 38058 Grenoble Cedex 9, France

kevin.parisot@gipsa-lab.grenoble-inp.fr, alan.chauvin@univ-grenoble-alpes.fr
ronald.phlypo@gipsa-lab.grenoble-inp.fr, steeve.zozor@gipsa-lab.grenoble-inp.fr

Résumé – La multi-stabilité visuelle est un phénomène dans lequel la perception varie au cours du temps tandis que le stimulus reste stationnaire dans le monde physique. Le système perceptuel bascule alors entre plusieurs états quasi-stables appelés percepts et il est possible d'utiliser ces stimuli pour étudier comment le cerveau choisi une perception face à l'ambiguïté. Développer des méthodes expérimentales adaptatives permet de compenser les fortes variabilités inter-sujet observées. Dans ce papier, nous présentons une méthode utilisant une modélisation probabiliste de l'observateur afin d'établir un protocole expérimental adaptatif. Les paramètres du modèle sont estimés par maximum de vraisemblance et le point d'ambiguïté maximale est obtenu chez les observateurs.

Abstract – Visual multi-stability is a phenomenon whereby perception varies while the stimulus remains stationary in the physical world. The perceptual system is said to oscillate between multiple states called percepts, and the phenomenon can be used to study how perceptual decision making in the brain occurs. High inter-individual variability in observers is a property that motivates the development of adaptive experimental methods. In this paper, we present a method that uses a probabilistic model of the observer in order to implement an adaptive experimental protocol. The model's parameters are estimated using maximum likelihood and the point of maximal ambiguity of percepts is obtained for each participant.

1 Introduction

La perception visuelle est étudiée depuis presque deux siècles en utilisant des stimuli dits multi-stables dans lesquels, la perception des observateurs change alors que stimulus reste stationnaire physiquement. Une variété de stimuli visuels engendrent ce phénomène, qui peut être utilisé pour étudier le traitement inférentiel de l'information sensorielle par le cerveau, dans sa reconstruction d'un monde perceptuel [1, 2]. Parmi les stimuli connus nous retenons le *plaid* en mouvement. Celui-ci présente deux grilles inclinées de manière symétrique parcourant un trou d'ouverture (voir centre de Fig. 1). Ce stimulus permet d'induire une tri-stabilité avec les percepts suivants :

— *cohérence* : une grille unique se déplace dans une seule direction,
— *transparence gauche/droite* : deux grilles qui se croisent et la grille qui semble la plus proche est celle qui se déplace vers la gauche/droite.

Notons la variable aléatoire X qui a pour ensemble de réalisations les différents percepts, i.e., $X \in \{c, g, d\}$ avec les percepts de cohérence (c), de transparence gauche (g) et de transparence droite (d), tous mutuellement incompatibles. Notre objectif est alors d'introduire un modèle probabiliste pour X et d'en estimer ses paramètres libres. Ceci permettra de prédire

FIGURE 1 – Exemples de *plaids* générés avec des valeurs de transparence : $\alpha_d = [0, 0.25, 0.5, 0.75, 1]$ et $\alpha_g = 1 - \alpha_d$

les liens entre les caractéristiques physiques du *plaid* [3, 4] et les probabilités des percepts pour chaque observateur. Effectivement, les probabilités d'observer les percepts peuvent être biaisées en manipulant les paramètres de présentation du *plaid*. De plus, ces probabilités montrent une grande variabilité inter-sujet [5] qui doit être contrôlée par l'expérimentateur avant toute analyse.

En vue du grand nombre de caractéristiques du *plaid*, il est tout d'abord nécessaire d'isoler un nombre restreint de paramètres qui permettent de manipuler efficacement la dynamique de perception. Nous avons ainsi cherché à vérifier s'il est possible d'influencer la dynamique de perception des sujets en contrôlant les paramètres de transparence des grilles du *plaid* en maintenant tous les autres paramètres constants par ailleurs.

Avant tout, il est nécessaire de vérifier la réalité terrain des percepts, ce que nous avons fait en concevant une première expérience d'identification des percepts.

2 Identification des percepts

La transparence des deux grilles (voir Fig. 1 gauche et droite) dans le stimulus peut être contrôlée par deux paramètres, correspondants à la transparence/opacité de chacune (un niveau 0 correspond à une transparence totale et un niveau 1 une opacité totale). Nous dénommons ces paramètres α_d et α_g qui correspondent au niveau de transparence de la grille en mouvement vers la droite et la gauche, respectivement [1]. Notre hypothèse est qu'il est tout à fait possible de contrôler les perceptions du stimulus en utilisant uniquement les transparences α_d et α_g ainsi faisant varier le rapport de luminances des deux grilles sans modifier la luminance totale du stimulus. En rendant les deux grilles opaques, le percept de cohérence serait dominant.

Dans cette expérience, nous avons donc exploré comment les sujets choisissent leur premier percept, connu pour avoir des propriétés particulières [3], en variant les paramètres $\boldsymbol{\alpha} = (\alpha_g, \alpha_d) \in [0,1]^2$. Nous cherchons à caractériser l'influence de ceux-ci vis-à-vis de la direction du mouvement perçu. Notons que nous formulons le percept en termes de direction du mouvement de la grille la plus proche selon l'observateur.

Premier protocole : vérification de l'influence d'alpha

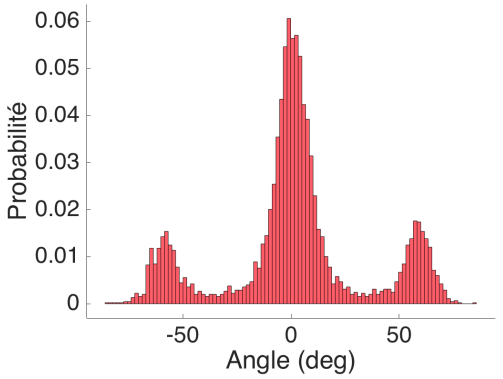
Le protocole expérimental consiste à générer des échantillons indépendants de $\boldsymbol{\alpha} = (\alpha_g, \alpha_d)$ distribués uniformément sur $[0,1]^2$ qui donne suite au stimulus présenté jusqu'à 5 secondes aux sujets pendant un essai (9 participants ; 10 blocs de 50 essais). Les sujets ont reçu l'instruction de répondre en cliquant avec la souris dans une zone de réponse, en périphérie du stimulus, pour indiquer la direction de fuite du mouvement du premier percept. L'angle relatif des grilles était constant, cependant un l'angle du percept cohérent (vers le haut pour les plaids dans la Fig. 1) était tiré aléatoirement, uniformément sur 360 degrés pour contrer tout biais de direction. Le contraste, i.e., la luminance moyenne, du stimulus a été tenu fixe.

Les réponses compilées (Fig. 2) montrent clairement l'existence de trois percepts liés aux directions du mouvement et sont facilement identifiables en fonction de $\boldsymbol{\alpha}$. Il est aussi notable que l'on peut réduire le couple de paramètres de transparence en suivant l'axe $\alpha_g + \alpha_d = 1$ (illustrée sur la Fig. 1) car la variabilité semble réduite le long de l'axe qui lui est orthogonal. Dans ce qui suit, nous dénommerons l'unique paramètre de présentation $\alpha = \alpha_d$ afin d'en simplifier la notation.

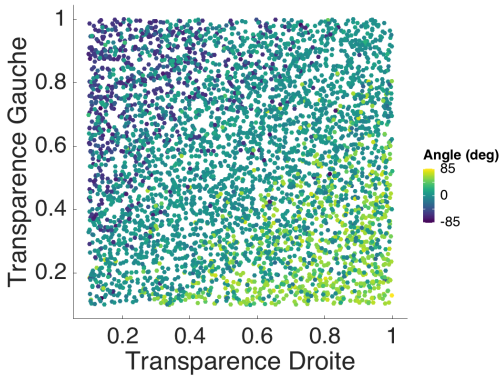
3 De la transparence à l'ambiguïté

En réduisant les paramètres de contrôle du stimulus à la transparence de la grille de droite α, on peut observer et mesurer le temps qu'un sujet passe dans un percept (t_d, t_g, t_c) et le

1. Remarquons que le déplacement en soi est ambigu car l'occlusion circulaire ne permet pas de différencier si la grille bouge vers le haut ou la droite, la perception choisie est alors souvent l'interprétation la plus simple d'un front d'une onde plane qui se propage orthogonalement aux transitions entre niveaux de gris.



(a) Angles reportés par les sujets.



(b) Angles perçus vs. (α_g, α_d).

FIGURE 2 – (a) Histogrammes des réponses sujets.Les angles ont été recalés par rapport à l'angle du percept cohérent c. (b) Angles des réponses souris relatives à l'angle du percept cohérent en fonction des transparences.

normaliser vis-à-vis du temps d'observation total, ce que nous traduisons en probabilité par marginalisation par rapport à t $\hat{p}(X = d|\alpha) = \frac{t_d}{t_d+t_g+t_c}$. On obtient ainsi une probabilité empirique pour le percept de *transparence droite* $p(X = d|\alpha)$ qui se veut d'être une fonction croissante car à $\alpha = \alpha_d = 1$ nous voyons que la grille dont le mouvement est vers la droite (celle de gauche étant complètement transparente). En faisant l'hypothèse de *symétrie* gauche/droite, i.e. courbes symétriques par rapport à $\alpha = 0.5$, on obtient celle du percept opposé :

$$p(X = g|\alpha) = p(X = d|1 - \alpha) \ . \tag{1}$$

Enfin la probabilité du percept de cohérence peut être calculée par *complémentarité* aux deux autres percepts en supposant une tri-stabilité du stimulus qui se base sur l'exclusivité mutuelle des percepts :

$$p(X = c|\alpha) + p(X = d|\alpha) + p(X = g|\alpha) = 1 \ . \tag{2}$$

Il est donc suffisant de définir une courbe paramétrique $p(X = d|\alpha)$ pour en déduire l'ensemble des probabilités par *symétrie* et *complémentarité*.

Le modèle que l'on utilise est la fonction sigmoïde, classique

en psychométrie :

$$p(X = d|\alpha, \boldsymbol{\theta}) = \frac{1}{1 + \mathrm{e}^{-\theta_1(\alpha - \frac{1}{2}) + \theta_2}} \ , \qquad (3)$$

où nous avons ajouté les paramètres $\boldsymbol{\theta} = (\theta_1, \theta_2)$ pour une régression en fonction du sujet (nous explicitons pas la dépendance sujet dans l'équation afin de ne pas surcharger les notations).

Il suit de la *symétrie* (1) et de la *complémentarité* (2) que $p(X = d|\alpha = 0,5, \boldsymbol{\theta}) \leq \frac{1}{2}$ et donc $\theta_2 \geq 0$. De plus, la fonction $p(X = d|\alpha, \boldsymbol{\theta})$ est croissante si et seulement si $\theta_1 \geq 0$.

Afin de déterminer les paramètres optimaux nous utilisons ici l'estimateur de maximum de vraisemblance $\widehat{\boldsymbol{\theta}}_{\mathrm{ML}}$. Ainsi, nous avons un modèle d'observateur psychophysique et une méthode d'estimation qui peuvent être couplés à une approche expérimentale qui permet un échantillonnage efficace.

Deuxième protocole : de la transparence à l'ambiguïté

Les études sur le *plaid* demandent que les sujets passent des expériences de visionnage continue, afin d'obtenir des données qui reflètent l'évolution temporelle de la perception face à un stimulus ambigu. Le premier percept est un cas particulier, mais peut être exploité pour gagner en efficacité sur la procédure d'estimation de $\boldsymbol{\theta}$. Ainsi, nous avons cherché à estimer la dépendance individuelle des participants vis-à-vis de α pour la tester dans un seconde temps en mettant en place les phases suivantes :

Phase 1 échantillonnage uniforme de α : 200 essais courts de mesure du premier percept,

Phase 2 échantillonnage normale tronqué ($\sigma = 0,05$) de α autour des points d'ambiguïté maximale α_{amb} (équiprobabilité des percepts), estimés en ligne, sur 10 essais de présentation de 2 minutes.

Contrairement à la première expérience, ici, l'orientation du mouvement n'était pas tirée aléatoirement, mais maintenue de sorte que la cohérence soit perçue comme un mouvement d'une grille vers le haut de l'écran. Les sujets devaient répondre à l'aide des touches du claviers correspondantes.

À la recherche d'une équi-probabilité

La première phase nous a permis d'explorer uniformément l'espace α et d'avoir une première estimation des paramètres $\boldsymbol{\theta}$. Dans la deuxième phase, il a fallu, à partir du modèle estimé, calculer le(s) point(s) d'ambiguïté maximale tri-stable α_{amb}. Il n'est cependant pas possible d'avoir un unique point d'équi-probabilité dans l'espace choisi, sauf si $p(X = d|\alpha = 0,5; \boldsymbol{\theta}) = \frac{1}{3}$ et donc $\theta_2 = \ln(2)$. Par contre, il est possible pour un ensemble $\boldsymbol{\theta}$ d'obtenir deux points symétriques tels qu'avec des probabilités $\pi(\alpha) = \pi(1 - \alpha) = \frac{1}{2}$ à chaque essai on obtient à long terme :

$$\begin{aligned} & \forall x \in \{d, g, c\} \ : \ p(X = x|\boldsymbol{\theta}) = \frac{1}{3} \\ = \ & p(X = x|\alpha, \boldsymbol{\theta})\pi(\alpha) + p(X = x|1 - \alpha, \boldsymbol{\theta})\pi(1 - \alpha) \ . \end{aligned}$$

Avec la symétrie on trouve que α satisfait $p(X = c|\alpha, \boldsymbol{\theta}) = 1/3$. On dérive ainsi l'expression suivante qui permet d'identifier l'ambiguïté maximale α_{amb}

$$\alpha_{amb} = \frac{1}{2} \pm \frac{\cosh^{-1}\left(\mathrm{e}^{\theta_2} - 2\mathrm{e}^{-\theta_2}\right)}{\theta_1} \ , \qquad (4)$$

ce qui admet une solution uniquement si $\mathrm{e}^{\theta_2} - 2\mathrm{e}^{-\theta_2} \geq 1$ ou $\theta_2 \geq \ln(2)$.

Résultats

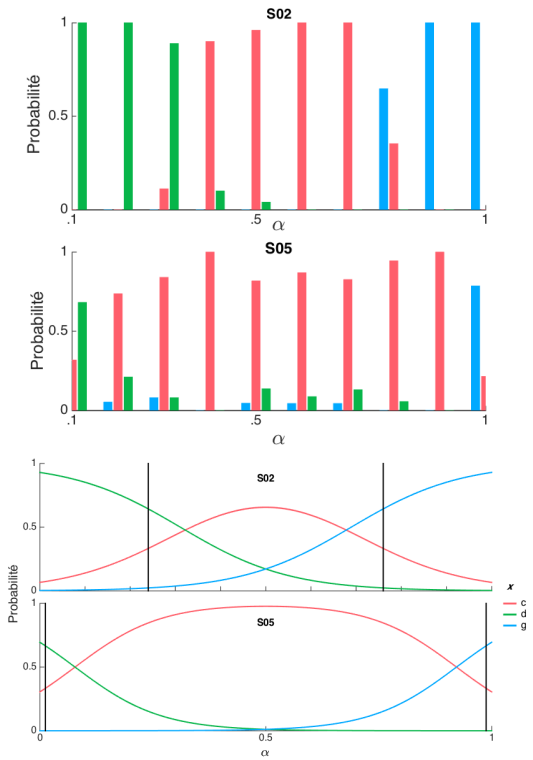
Les résultats présentés dans les Fig. 3 montrent la dépendance de α_{amb} des sujets ($N = 12$ sujets). Les histogrammes représentent les réponses données lors de la phase 1, et les courbes les modèles estimés à la fin de la phase 1. Cette expérience montre que l'on peut contrôler la variabilité intersujet de cette manière en calibrant le stimulus. La Fig 3a montre que les courbes psychophysiques basées sur la fonction $p(X = x|\alpha, \widehat{\boldsymbol{\theta}})$ conviennent particulièrement bien pour la modélisation des données dans le cas d'un sujet comme le S02. De plus, Fig. 3b montre que l'estimateur par maximum de vraisemblance reste stable lors de la phase 2.

4 Discussion

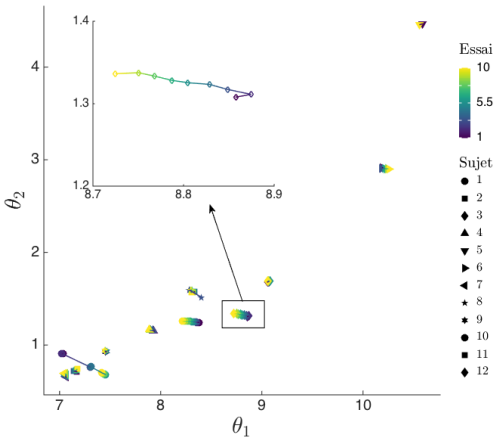
Ces modèles et expériences pilotes sont des outils qui peuvent être appliqués afin de calibrer les paramètres du stimulus du *plaid* pour chaque participant. Nous avons développé ces techniques afin de les intégrer dans une expérience plus complexe qui a pour objectif de montrer des effets de causalité et les corrélations entre le système oculomoteur et perceptif. Comme ce dernier n'est pas directement accessible par l'expérience, il est donc nécessaire d'avoir un contrôle sur le système perceptif en utilisant uniquement l'information visuelle contenue dans le stimulus. Ainsi, sous notre modèle il est désormais possible de prédire (les probabilités de) la perception en fonction de l'ambiguïté.

Dans ce travail, plusieurs choix restent à étudier pour s'assurer de l'optimalité du processus d'estimation et d'application du modèle pour ce stimulus. Il est possible de varier le choix de la fonction sigmoïde pour correspondre d'avantage aux observations. Par exemple, lorsque formulé comme l'équation (3), on obtient une contrainte sur la position du paramètre θ_2 qui permet de s'assurer que le modèle ne puisse pas estimer un seuil de dominance du percept de la *transparence droite* quand la grille droite est plus transparente que celle de gauche. *A posteriori*, après analyse des comportements des participants, l'écriture de la fonction sigmoïde peut être modifiée pour qu'il soit possible que le modèle génère trois distributions uniformes dans le cas de sujets qui répondent au hasard.

L'échantillonnage des "prochains" (α_d, α_g) avec estimation en ligne a été, ici, fait selon la procédure décrite ci-dessus, cependant, il est aussi possible de tirer de façon plus large dans l'intervalle qui est compris entre les points d'inflexions qui mène aux plateaux de saturation de $p(X = d|\alpha, \boldsymbol{\theta})$ et



(a) Phase 1 : histogrammes et modèles des réponses.



(b) Phase 2 : évolution de $\boldsymbol{\theta}$ au cours de l'échantillonnage.

FIGURE 3 – (a) Histogrammes des probabilités de réponses de la phase 1 de l'expérience 2 en fonction de α pour S02 et S05. Ci-dessous, le modèle avec les paramètres inférés à la fin de la phase 1. Les barres noires montrent les points α_{amb} calculés pour chaque sujet. (b) Évolution des paramètres $\boldsymbol{\theta}$ au cours de l'échantillonnage dans la phase 2 pour tous les sujets. L'évolution de l'estimation des paramètres du modèle de S03 est agrandi.

$p(X = g|\alpha, \boldsymbol{\theta})$. Cet intervalle contient théoriquement le maximum d'*information* sur la dynamique interne des sujets et donc, l'estimation des paramètres du modèle complet.

Les approches expérimentales adaptatives sont développées depuis quelques décennies maintenant avec des variétés de procédures proposées comme QUEST ou ZEST [6, 7], des méthodes notamment utilisées en psychophysique. Il est même possible de développer une approche plus générale en se basant sur la méthode de réponse expérimentale obtenue et à partir des concepts d'inférence de paramètres de systèmes bayésiens [7]. Ces méthodes généralisées sont susceptibles de s'adapter à la distribution des paramètres variant du stimulus, plutôt qu'au processus de perception du sujet. Un travail à poursuivre serait de mettre en place ces méthodes sur ce type de données et de les comparer à notre approche.

5 Conclusion

Nous avons présenté une méthode de manipulation de l'état perceptif de participants dans un phénomène de multi-stabilité perceptuelle induit par le stimulus du *plaid* en mouvement. Il est possible de réduire la manipulation à la transparence d'une grille, d'estimer un modèle probabiliste de chaque individu et de mettre en place un protocole expérimental adaptatif. Cette approche prend en compte des contraintes théoriques et expérimentales permettant une inférence rapide des paramètres, afin de calibrer le stimulus à la perception des sujets. Ces méthodes ouvrent la possibilité de contrôler efficacement le niveau d'ambiguïté perçue et de tester des hypothèses d'influences de systèmes (par exemple moteur, émotionnel, etc) en interaction avec celui-ci.

Références

[1] D. A. Leopold and N. K. Logothetis, "Multistable phenomena : changing views in perception," *Trends in cognitive sciences*, vol. 3, no. 7, pp. 254–264, 1999.

[2] J. A. S. Kelso, "Multistability and metastability : understanding dynamic coordination in the brain," *Philosophical Transactions of the Royal Society B : Biological Sciences*, vol. 367, pp. 906–918, feb 2012.

[3] J.-M. Hupé and N. Rubin, "The dynamics of bi-stable alternation in ambiguous motion displays : a fresh look at plaids," *Vision research*, vol. 43, no. 5, pp. 531–548, 2003.

[4] G. Huguet, J. Rinzel, and J.-M. Hupé, "Noise and adaptation in multistable perception : Noise drives when to switch, adaptation determines percept choice," *Journal of Vision*, vol. 14, pp. 19–19, 03 2014.

[5] J.-L. Schwartz, N. Grimault, J.-M. Hupé, B. C. Moore, and D. Pressnitzer, "Multistability in perception : binding sensory modalities, an overview," 2012.

[6] A. B. Watson, "Quest+ : A general multidimensional bayesian adaptive psychometric method," *Journal of vision*, vol. 17, no. 3, pp. 10–10, 2017.

[7] J. H. Bak and J. W. Pillow, "Adaptive stimulus selection for multi-alternative psychometric functions with lapses," *Journal of vision*, vol. 18, no. 12, pp. 4–4, 2018.