

# Peut-on mesurer la beauté d'une photo ?

Henri Maître

LTCI, Télécom ParisTech, Université Paris-Saclay

46 rue Barrault - 75013 Paris

henri.maitre@telecom-paristech.fr

## Résumé

*Si la notion de qualité d'image est un sujet d'étude des traiteurs d'images depuis plus de quarante ans, les notions liées à l'esthétique des photographies ne sont apparues que depuis une dizaine d'années dans cette communauté. Mais avec l'automatisation des fonctions d'archivage et de recherche dans les très vastes bases de données numériques, les études qui sont consacrées à ce thème se multiplient aujourd'hui, mettant à profit le développement des techniques d'apprentissage ainsi que la prolifération des sites spécialisés dans l'archivage des photos sur internet. Avant d'examiner l'avancement de ces méthodes numériques, nous nous attardons sur d'autres approches de l'esthétique : en philosophie, en neuro-biologie, en psychologie expérimentale et en sociologie, de façon à voir quels éclairages elles proposent au chercheur. Nous soulignons la faiblesse des consensus actuels sur ce problème difficile de l'esthétique et l'importance des débats encore en cours.*

*Nous apportons ensuite un regard synthétique sur les travaux menés dans la communauté de reconnaissance des formes et d'intelligence artificielle. Nous comparons les résultats présentés et nous nous livrons à un examen critique des démarches adoptées.*

## Mots Clef

Esthétique, beauté, image, apprentissage automatique, réseaux neuro-mimétiques

## Abstract

*Image quality has been introduced in image processing community studies at least 40 years ago, in contrast to notions linked with aesthetical qualities of photographs which only appeared in the computer science literature in the last decade. However, because of the evolution of Information and Communication Technologies and its consequences on business and societal practices (storage on very large and distributed data bases, archival and retrieval functions, automatic learning processes) and because of the flourishing of photographic exchange sites, the domain received a growing attention in the last decade. Digital aesthetic image assessment is now rich of some tens of papers and "Image Beauty" is more and more chosen as a major criterion to select images in retrieval operations. Is this*

*criterion mimicking human judgement relevant ? Is it correctly translated into algorithms ?*

*How reliable may be approaches which are only moderately supported by human expertise ? Which benefit is hired from 20 centuries of literature on aesthetics, beauty and art, in Philosophy, in Sociology and in experimental Psychology as well ? Which relations are weaved with the ongoing studies in Physiology and in Neuro-Biology that intend to analyze how our brain is working, some of them directed towards explaining our aesthetical judgement, now recognized under the name of "neuro-aesthetics" ?*

*Here are some questions we try to address.*

## Keywords

Aesthetics, beauty, image, machine learning, CNN.

## 1 Introduction

Depuis trois ans, les communications se multiplient dans la communauté de reconnaissance des formes et d'intelligence artificielle (IA), proposant de fournir une évaluation automatique de la valeur esthétique d'une image [1, 2, 3, 4, 5, 6, 7]. Ces travaux font appel à des réseaux de neurones profonds (RNP). Ils font suite à une poignée de travaux, démarrés au début du siècle, qui abordaient ce problème à l'aide de techniques plus classiques d'apprentissage : détection de primitives choisies par l'utilisateur (*hand-crafted*) et classifieurs de divers types [8, 9, 10, 11, 12]. Bien sûr, les techniques par RNP ont rapidement surclassé les méthodes plus traditionnelles comme elles l'ont fait dans plusieurs autres domaines de la reconnaissance des formes. Ces travaux n'ont été rendus possibles que grâce à un certain nombre de progrès que nous allons examiner. Ils prennent place également dans un contexte scientifique qui permet d'aborder ce problème par des voies très différentes, en particulier par des démarches de neurobiologie ainsi que par des expérimentations de psychologie sociale. Quel est le rôle de ces travaux ? Quelle influence peuvent-ils avoir sur l'évolution des techniques d'IA ?

## 2 Mesurer la beauté

Les premiers travaux qui ont proposé une mesure mathématique de la beauté sont dus à Charles Henry [13], mais c'est le mathématicien Birkhoff qui a proposé la première formulation opérationnelle [14]. Cette formulation, inspirée par 20 siècles de littérature philosophique sur

l'esthétique dans les arts en particulier visuels, était bâtie sur les notions d'ordre et de simplicité à une époque où ces deux termes n'avaient guère de sens en mathématiques. Elle s'est enrichie au cours du siècle dernier par les apports successifs de la Gestalt, de la théorie de l'information, de la morphologie mathématique et de la théorie de la complexité pour aboutir à des expressions algorithmiques et algébriques [15, 16, 17] qui, malgré leurs résultats intéressants, sont restées très confidentielles.

Les techniques fondées sur l'apprentissage automatique qui ont débuté avec ce siècle ont fait table rase de ces travaux. Elles ont au contraire exploité opportunément un contexte scientifique nouveau : l'abondance des images accessibles sur Internet, la disponibilité de nombreuses sources d'expertise à travers les réseaux sociaux spécialisés ou grand-public, et enfin l'émergence de techniques statistiques puissantes pour apprendre des règles de classifications et la possibilité de les étendre à des grands ensembles inconnus. Puis la diffusion des techniques de RNP, exploitant successivement des filtres convolutionnels puis des décisions totalement connectées, a consacré les approches "aveugles" pour lesquelles l'expertise humaine se réduit souvent à la constitution des bases de données indexées permettant l'apprentissage.

Cette rupture complète dans les paradigmes à la base de la démarche esthétique mérite que l'on prenne le temps d'examiner les conséquences que l'on peut attendre de ces choix et de mettre en parallèle les attentes des utilisateurs et les résultats potentiels des méthodes employées.

## 2.1 Les approches philosophiques

Il n'est guère envisageable de traiter en quelques lignes les très nombreux ouvrages qui ont progressivement fait de l'esthétique une partie propre et reconnue de la philosophie. Les plus grands noms s'y sont essayés [18, 19, 20, 21]. On opposera schématiquement l'école "objectiviste", issue des philosophes grecs, (qui défend une idée universelle du Beau, attachée à l'objet ou à la personne qu'il qualifie, idée qui est partagée par tous et en tout lieu), à l'école subjectiviste, née des philosophes des Lumières (qui rapporte le Beau à l'expérience individuelle et ses contingences expérimentales). Les grands courants de pensée qui, depuis la psychanalyse ont parcouru la philosophie ce dernier siècle ont transporté ce problème en des termes plus modernes : le Beau est-il si unanimement perçu parce qu'il active sélectivement des sensations physiologiques universelles ou n'est-il que le résultat d'influences biochimiques et environnementales conjoncturelles et individuelles [22, 23, 24, 25] ? Loin d'être tranché, le débat rebondi perpétuellement, profitant des éclairages nouveaux de la science.

Et ce débat impacte directement notre projet de mener à bien un "mesureur de beauté". Faut-il copier les "beaux objets" et en découvrir les canons ou faut-il comprendre les ressorts de l'émotion et de la conscience afin de leur livrer une matière sélectionnée pour les satisfaire ?

Parmi les résultats importants de ces travaux philosophiques, notons des contributions marginales, mais indispensables à notre propos, par exemple pour distinguer clairement la part de l'esthétique dans l'Art (en particulier contemporain) [26, 27, 28], ainsi que pour démêler des sources multiples d'intérêt qui pourraient masquer le rôle de l'esthétique dans l'attrait d'une œuvre [29, 30, 31].

## 2.2 Les approches de la neuro-esthétique

Si notre monde scientifique s'est livré avec passion aux réseaux neuro-mimétiques, il s'était livré, depuis 30 ans et avec le même enthousiasme, aux perspectives de l'imagerie cérébrale. L'imagerie par résonance magnétique, puis l'imagerie par résonance magnétique fonctionnelle (IRMf) ont apporté des outils exceptionnels de compréhension du fonctionnement de notre cerveau. Très tôt utilisée pour comprendre les règles mystérieuses de notre jugement artistique, l'IRMf a ainsi donné naissance à une branche distincte de la neuro-biologie, qui se reconnaît sous le nom de neuro-esthétique [32]. La littérature en contient plus de 3000 publications, la plupart consacrée aux arts visuels.

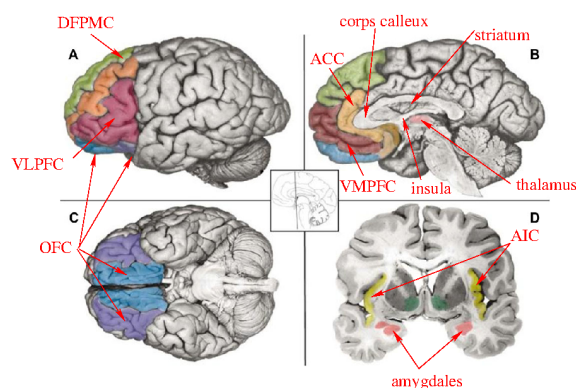


FIGURE 1 – Les aires concernées lors de l'expression d'émotions, d'après [33]. Le cortex orbito-frontal (OFC) est la zone en violet des vues A et C. Le cortex insulaire (AIC) est en jaune sur la vue D. Il est relié de façon très étroite aux aires latérales de l'OFC. Les circuits viscéro-moteurs comportent la zone ventrale du cortex ventro-médian préfrontal (VMPMC) (en bleu sur A, B et C). Le VMPMC est étroitement lié à l'amygdale (en rose dans la vue D). Le cortex cingulaire antérieur (ACC) est en beige sur B, en avant du corps calleux. La zone dorsale du cortex préfrontal est associée au contrôle des états mentaux : partie dorsale du VMPMC, zone frontale (en marron), cortex dorso-médian préfrontal (DMPFC) (en vert en A et B). Le cortex ventrolatéral préfrontal (VLPFC) est en rouge sur la vue A. Le thalamus est en rose au centre de la vue B et le gyrus frontal moyen est en orange.

Que nous apprend la neuro-esthétique ? Beaucoup plus que ce que l'on peut résumer en quelques lignes, bien sûr, mais on trouvera dans [34] une très bonne synthèse de ces travaux. Tout d'abord, elle nous permet d'écarter l'idée quelque temps envisagée, d'"aires hédoniques", spécialisées dans le traitement du beau ; les aires cérébrales

mobilisées apparaissent, au contraire, toutes déjà engagées dans des traitements cérébraux différents, opportunément réutilisés. Quelles sont ces zones ? Elles sont brièvement présentées sur la figure 1 et rappelées ci-dessous :

- les aires visuelles, les zones occipitales et latérales inférieures, le cortex insulaire, le lobule pariétal supérieur sont bien sûr actifs par la tâche même de vision, mais aussi par l’extraction du contenu selon des aires spécifiques : formes, couleurs, mouvements, visages ;
- le cortex orbito-frontal en charge de l’évaluation, régulateur des “récompenses” et du plaisir, est une pièce importante qui semble contrôler notre décision ;
- le cortex insulaire, qui contrôle nos émotions, est lui aussi une pièce importante et quasiment toujours impliquée lors de l’observation d’une œuvre d’art ;
- les zones engagées dans les opérations de cognition (amygdale) et de mémoire (aires pariétales médianes, lobe préfrontal) sont souvent sollicitées ;
- des aires en charge du contrôle prémoteur sont convoquées de façon spécifique dans des situations de forte empathie et d’*embodiment* (cortex ventral prémoteur, lobes temporaux, hippocampe).

On trouve dans [35, 36, 37, 33] de nombreuses précisions sur les expériences conduites, les résultats obtenus ainsi que certaines conclusions que l’on peut en tirer. Dans [38] est présentée une vision synthétique de nos connaissances sur ce sujet.

On ne peut cependant manquer de souligner combien l’IRMf, en l’état de son développement, est insuffisante à comprendre les mécanismes réellement mis en œuvre : les temps de réponse des instruments, la nécessité de moyenniser sur les expériences et sur les individus altèrent les capacités déductives de ces travaux. Il est en particulier pratiquement impossible de remonter à la chronologie entre stimuli par les voies visuelles et activation des aires supérieures, conditions indispensables à une véritable explication par causalité. C’est à ce prix pourtant que l’on éclairera le débat entre “objectivistes” et “subjectivistes”. Notons également que ces travaux se heurtent à des critiques théoriques plus profondes [39, 40].

### 2.3 Psychologie expérimentale, psycho-sociologie et photographie

Une autre source importante d’information sur l’esthétique provient de la littérature consacrée à la photographie, et en particulier des recommandations issues des manuels de photographie et des ouvrages de photographie artistique. La pertinence de ces recommandations peut se mesurer de diverses façons : leur fréquence dans les ouvrages (c’est par exemple le cas de la règle du tiers), la popularité de leur auteur, voire sa cote sur le marché de l’art (qui sollicitent alors la démarche sociologique [41]), mais aussi les vérifications qui peuvent être conduites à l’aide des tests de la psychologie expérimentale ainsi que les vérifications statistiques sur des corpus.

réf.	base de données	primitives et classifieur
[8]	DPChallenge	attributs génériques - Bayes, AdaBoost
[53]	DPChallenge	primitives photo - Bayes, SVM, AdaBoost
[9]	Photo.net	primitives photo - SVM
[10]	CUHK	attributs génériques - PCA+FV+SVM
[54]	DPChallenge	primitives haut niveau - SVM
[55]	Photo.net	compos. texture,color - Acquine+SIMPLcity
[56]	CUHK	primitives photo - SVM par catégorie
[57]	DPChallenge	opinions internautes - SV-ε
[58]	AVA	HSV ou Munsell - LDA + Lasso
[11]	AVA	photo + ss-photo - RNP - SCNN + RD-CNN
[12]	BEAUTY	primitives photo - PLSR
[59]	AADB+AVA	image réduite + catégorie - RNP “siamois”
[3]	AVA	pyramide + info class. - RNP multitâche MNA
[60]	AVA	image + 3 catégories - A&C RNP
[4]	AVA	histo. de notes. - Brain-inspired RNP
[61]	Redi	apprentissage images web - RNP
[5]	AVA	esthétique + catégorie - RNP multitâche
[6]	AVA	composition + ss-images - A-Lamp, multitâche
[7]	AVA	histo. de notes.+catégorie - RNP multitâche

TABLE 1 – Les systèmes d’apprentissage de l’esthétique, leurs bases d’apprentissage et les principes de classification.

Disons tout de suite que les travaux sur ce sujet sont très décevants : bien peu de règles résistent aux vérifications objectives. On est tout d’abord amené à écarter des critères classiquement importants pour définir la “qualité” de l’image (résolution, piqué, rapport signal à bruit, contours) [42, 43], car qualité et beauté évoluent dans des espaces subjectivement différents. On élimine également bon nombre de “règles de composition” : règle du tiers, nombre d’Or, poids du centre, symétries et orientations privilégiées, etc. [44, 45, 46].

Les règles sur la répartition des ombres et lumières semblant valider les décroissances régulières en  $1/f^2$  du spectre de densité de puissance sont assez bien vérifiées [47], tandis que les lois sur l’histogramme des niveaux de gris se réduisent à des préférences assez difficiles à imposer sur les moments d’ordre 3 [48]. Enfin les préférences sur la palette chromatique, qui semblaient bien ancrées avec les théories bien établies de Moon et Spencer et de Matsuda, se heurtent à des réfutations frontales [49, 50].

L’universalité des critères esthétiques est donc souvent battue en brèche lorsqu’on se réfère à ces témoignages. Des études sur l’examen du parcours de l’œil de l’observateur au cours de l’examen de la photo [51, 52] confirment que les règles de parcours d’un document sont très dépendantes du bagage culturel de l’observateur.

## 3 Les approches par apprentissage

Nous l’avons dit, les premiers travaux de mesure de l’esthétique étaient algébriques (ils ne faisaient pas appel aux techniques d’apprentissage automatique) et n’ont pas entraîné une très large adhésion de la communauté. Les travaux suivants sont partis sur des bases très différentes. Ils sont présentés dans le tableau 1.

### 3.1 Handcrafted features et classification

Les premières publications faisant état de travaux destinées à apprendre automatiquement les règles de l'esthétique des images remontent à 2006 [62, 8]. Ces travaux s'appuient sur la disponibilité de nombreuses bases de données d'images sur Internet, bases de données qui sont souvent accompagnées d'informations permettant de juger de la beauté de chaque image. La nature de ces bases de données est très variable. Certaines sont de simples archives échangées entre particuliers (type Flickr), d'autres sont destinées aux amateurs, voire aux professionnels, de la photographie (comme Photo.net ou DPChallenge), d'autres enfin sont réservées aux scientifiques (Image CLEF, Beauty, AVA). Les annotations reflètent également des avis très divers, depuis les "like" du grand public, jusqu'aux avis d'experts de jurys photographiques, en passant par les consultations spécialisées. Le schéma de prin-

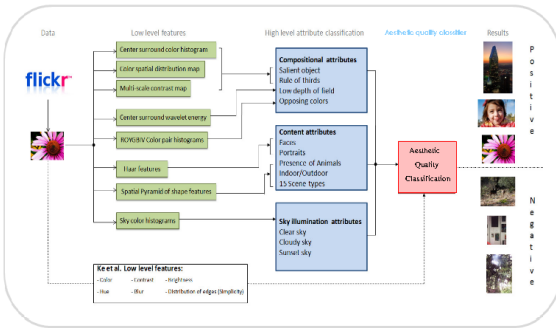


FIGURE 2 – Le schéma de calcul des primitives : dans ces travaux de S. Dhar et al. [54], les primitives sont calculées à deux niveaux : tout d'abord au niveau bas, comme dans la plupart des travaux sur l'esthétique, puis à un niveau d'abstraction plus élevé par combinaison des primitives de bas niveau. Toutes ces primitives sont utilisées par un classifieur unique.

cipe des méthodes à base de primitives est représenté sur la figure 2. Les divers travaux présentés dans ce cadre diffèrent par le choix des primitives et par le classifieur [55, 63, 12, 64]. Les primitives classiquement utilisées dans le domaine du multimedia [9, 65] ont été employées concurremment à des primitives plus spécialisées dans le domaine de l'esthétique [53, 54, 58] sans montrer de perte de performances. Les classifieurs utilisés s'appuyaient sur des techniques bayésiennes, des SVM ou des décisions sur graphes.

### 3.2 Réseaux de neurones profonds (RNP)

Dès leur apparition, les techniques à base de RNP ont montré des performances supérieures aux approches plus classiques. Les architectures adoptées sont celles que l'on retrouve dans tout le champ de la reconnaissance dans les images : des couches de convolutions suivies de couches totalement connectées.

Des évolutions sont cependant apparues pour adapter ces systèmes aux spécificités du problème (voir la figure 3) :

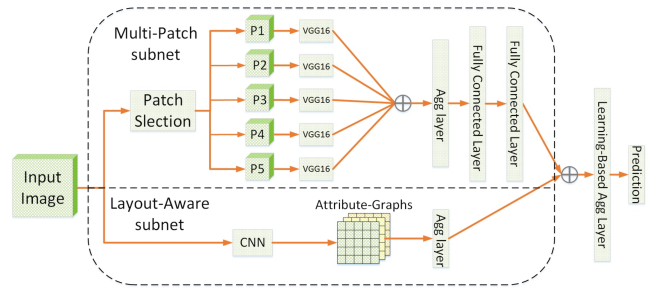


FIGURE 3 – Exemple d'architecture de RNP pour l'étude de l'esthétique : le système A-Lamp [6]. Cette architecture est constituée de 5 RNP (ici des architectures VGG16, travaillant en parallèle sur des sous-parties de l'image à pleine résolution et chargés d'évaluer certaines propriétés probablement importantes pour l'esthétique) et d'une voie en charge d'extraire un graphe d'attributs de l'image totale.

- plusieurs solutions ont été proposées pour permettre de traiter de très grandes images en conservant la fine structure des détails : présélection de fenêtres autour de points d'intérêt [6, 66], traitement en parallèle de fenêtres tirées aléatoirement [58], utilisation de structures hiérarchiques [67, 3], etc. Malgré ces travaux, la taille des couches d'entrée des RNP opérationnels constitue une limite aux travaux sur l'esthétique qui manipulent toujours des images de grande dimension ;
- la prise en compte d'informations annexes, très importantes dans le choix des critères à appliquer, a conduit à des réseaux à flux multiple [5, 6] qui exploitent diverses connaissances : le type d'image, le style de la photo, la classe de l'objet principal, etc. ;
- la reproduction de certains mécanismes cérébraux a amené à séparer l'architecture de traitement selon des voies différentes [55, 5] ou, parfois, en une succession de RNP : l'un en charge du bas niveau, un autre en charge des informations de haut niveau [4].

La mise en place de techniques à base de RNP a fait sensiblement évoluer les travaux menés sur l'esthétique des images.

Un premier élément de différenciation concerne le choix des bases de données. La nécessité de disposer de très vastes bases d'apprentissage a fait abandonner des travaux qui utilisaient des bases de données originales de quelques milliers d'images seulement. La communauté s'est ainsi concentrée sur la base AVA qui a le mérite de disposer d'images souvent très belles (elles sont issues du site spécialisé pour photographes DPChallenge) dotées de très nombreux avis sur chaque image. Cependant, pour la nécessité d'entraînement des réseaux, sa taille (elle est constituée de 250 000 images) est souvent insuffisante. On peut alors avoir recours à l'extension artificielle par manipulation d'images [4], mais plus généralement on utilise des réseaux pré-entraînés à l'aide des protocoles mis en place pour Image.net (voir par exemple [5]) pour des ap-

plications de reconnaissance de catégories.

Un second élément est la disparition quasi-complète (excepté dans les travaux [6]) des critères esthétiques pour la construction de l'architecture de RNP. Les travaux qui font appel à des informations extérieures à l'image utilisent surtout des données sur le type d'image : intérieur, portrait, sport, etc., données qui semblent cependant assez lointainement liées à la beauté de l'image.

### 3.3 Des performances en constant progrès

Les études d'esthétique sont très difficiles à évaluer. La base de données (AVA [11]) a été créée pour pallier ce problème. Pour cela, elle a été dotée d'un très large corpus d'annotations de qualité sur une échelle de 1 à 10 (près de 200 notes par photo). Cette base est souvent utilisée en séparant les échantillons en deux familles : les "belles images" (dont le score moyen est supérieur à  $5 + \delta$ ) et les "vilaines images" (dont le score moyen est inférieur à  $5 - \delta$ ). Le paramètre  $\delta$  permet à chacun de séparer les classes lors de l'apprentissage avec la rigueur qu'il souhaite. Il est souvent pris nul ou égal à 1. La figure 4 montre l'évolution des performances de reconnaissance au cours des dernières années pour  $\delta = 0$ . On observe une croissance régulière de ces performances.

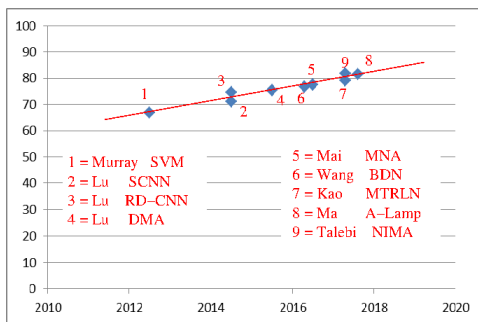


FIGURE 4 – Evolution des performances de reconnaissance des belles images de la base AVA avec une valeur  $\delta = 0$  au cours des 5 dernières années (exprimée en pourcents). Seuls les travaux de Murray [11] utilisent des primitives et un classifieur par SVM, tous les autres utilisent une classification par RNP).

## 4 Analyse des travaux

On peut aborder une analyse des travaux conduits par la communauté de traitement des images pour évaluer la beauté des photos sous plusieurs angles. Nous en choisissons ici quelques uns qui n'épuisent pas l'analyse.

### 4.1 Des propriétés peu exploitées

L'analyse d'une photo par les architectures classiques de RNP ont montré leur puissance pour reconnaître et localiser des objets, même déformés ou partiellement masqués. Il semble cependant que certaines propriétés importantes de l'évaluation esthétique nécessiterait de faire évoluer les architectures. Nous avons déjà signalé l'importance de pouvoir traiter de grandes images avec de nombreux détails

fin. Signalons également l'importance qu'il faudrait accorder à l'harmonie chromatique qui est indéniablement une composante importante de l'esthétique (les travaux de [66] sont à ce titre exemplaires). Il n'est pas évident que des architectures qui procèdent à des convolutions dans les premières couches respectent bien la subtilité des nuances. La construction interne de la photographie est elle-même un élément important de la qualité esthétique de l'œuvre (R. Diderot en faisait un argument majeur de sa démarche esthétique [68]). Reconnaissons que, si de nombreux travaux essaient d'en tenir compte, bien peu se donnent les moyens de le faire à travers les couches d'abord convolutionnelles puis totalement connectées du RNP. A notre connaissance seul les auteurs de [6] réservent une voie de traitement à cette structure.

### 4.2 Les critères : "belle" vs. "vilaine" image

Le critère binaire adopté par la communauté pour comparer les diverses approches a de nombreux mérites. Il peut être appliqué rapidement sur de très grandes bases ; il se transporte aisément d'une base à une autre ; il se prête à une vérification visuelle simple ; il offre une bonne solution à quelques problèmes que la société se pose : trier très vite de grandes archives pour en garder une quintessence, fournir des exemples attrayants pour des illustrations, assister un opérateur pour sa prise de vue, etc.

Il souffre cependant d'être exagérément simplifié. Il s'appuie sur le postulat que toute image relève de l'une ou l'autre catégorie, postulat dont on ne trouve nulle trace dans la littérature. Bien plus, il est très communément admis, aussi bien en philosophie qu'en neuro-biologie, que l'attribut de beauté n'a qu'une valence positive et pas d'équivalent à valence négative (qui serait nommé le laid), cette valence étant prise en charge par d'autres attributs comme "effrayant", "triste", "ennuyeux", "banal", "brouillon", etc.

Ainsi, la complexité des informations transmises à travers les notations pour chaque image de la base AVA est aujourd'hui insuffisamment prises en compte, même si certains travaux essaient de les exploiter [69, 7, 4]. Il serait cependant important de distinguer les dispersions de notes dues à l'hétérogénéité d'intérêt, d'attention, de culture, de motivation ... des experts, de ce qui relève des propriétés intrinsèques de la photo (ce que les auteurs de [69] attribuent à une "difficulté" inhérente d'interprétation).

### 4.3 Classer des images

Dès l'origine [62], de nombreux travaux s'étaient fixé pour objectif de classer des photos selon une échelle de beauté plus ou moins continue. Quoique de nombreux algorithmes fournissent une note entre 0 et 10, peu de travaux rendent compte de la qualité de ces notations [7] hormis pour affiner la décision binaire [60, 4]. L'évaluation d'un classement continu est aujourd'hui très difficile et nous semble cependant un enjeu majeur. Notons que dans [12] une classification en cinq niveaux permet d'affiner sensiblement la mesure. Notons surtout l'approche très originale de [70]

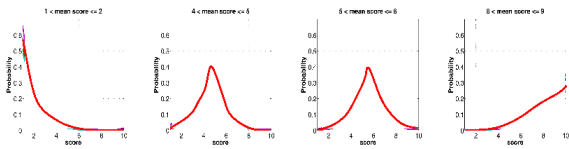


FIGURE 5 – Distribution moyenne des notes attribuées à chaque image au sein de la base AVA pour des valeurs moyennes croissant de gauche à droite : entre 1 et 2, entre 4 et 5, entre 5 et 6, entre 8 et 9 (d’après [11])

qui propose de comparer deux à deux les images de la base pour atteindre une évaluation relative.

#### 4.4 Quelle beauté ?

Les images utilisées pour les tests de performances représentent bien ce que l’on peut attendre d’images de qualité issues des réseaux sociaux. Les ”belles images“ sont incontestablement généralement supérieures aux images ”vilaines“. Si les qualités des images ”belles“ ne sont pas toujours évidentes, on constate qu’elles montrent rarement les défauts qui font écarter les images ”vilaines“ : mauvaise composition, mauvaise répartition chromatique, défaut de mise au point, etc.

Un observateur attentif et exigeant sera cependant assez souvent en désaccord avec les décisions faites par le système, même si ces décisions sont en conformité avec les jugements portés par la base d’expertise. Cela s’explique souvent soit parce qu’une ”belle “ image se révèle banale, soit, surtout, parce qu’une image de qualité a été classée ”vilaine“. On constate dans ce dernier cas que ce sont fréquemment des aspects originaux de l’image qui ont été ignorés. Les RNP privilégient des images ”dans la norme“, ce qui n’est guère en accord avec les recommandations des experts.

On peut regretter à ce propos qu’aucun système ne se soit confronté aux photos réputées remarquables des archives photographiques. Il y aurait certainement beaucoup à apprendre du dépouillement objectif de tels résultats.

#### 4.5 Quelle expertise ?

On aborde là un des points très sensibles de l’approche par RNP. L’importance de disposer d’une base de données de qualité assez vaste a été ressentie dès la mise en place des approches par *handcrafted features*, mais elle est devenue cruciale pour les approches neuromimétiques. La base AVA [11] a permis d’apporter une bonne réponse à cette demande. Au-delà de la collection d’images, AVA est forte de plusieurs informations attachées à chaque photo : les évaluations, le thème couvert par l’image (parmi plus de 900, issus des compétitions de la banque de photo DPChallenge d’origine), une annotation sémantique (parmi 66, issue des thèmes) et le style photographique (attribué par des photographes professionnels parmi 14).

Est-ce suffisant ? Ce n’est pas certain. Certes pour les ob-

jectivistes qui placent toute la beauté dans le seul objet, il y a dans AVA l’objet fidèlement reproduit et dans la note moyenne l’expression du consensus sur son appréciation. Il y a donc tous les éléments suffisants à entraîner une machine à reproduire le jugement humain, pour peu que l’on maîtrise les rouages de l’apprentissage artificiel.

Si l’on accorde à l’observateur une place plus importante, on est plus exigeant sur les informations qui seront nécessaires à l’évaluation. Sans adopter les positions extrêmes des subjectivistes qui attribuent la totale autorité sur le jugement aux humeurs de l’observateur, on peut attendre d’autres informations pour simuler un sentiment qui en appelle aux sensations d’une part (celles convoiées par le signal visuel de l’image), aux facultés mentales conscientes et inconscientes de l’observateur d’autre part et à son tempérament enfin. Que l’on puisse tirer ces informations de la base AVA est très douteux. Ainsi, dans [59], on a jugé nécessaire de construire une base (AADB), différente d’AVA, en gardant lors de l’évaluation la marque de l’évaluateur, pour que l’on puisse entraîner la machine à évaluer ”à la façon“ de X ou Y. Les auteurs indiquent qu’un tel choix permet d’obtenir une meilleure adéquation entre les classements obtenus par un même expert. Dans [12] on a porté attention au contexte culturel des experts utilisés pour construire la base BEAUTY. Seuls ont été retenus des internautes issus d’un petit nombre de pays à forte homogénéité culturelle et leurs avis ont ensuite été filtrés a posteriori pour écarter les points de vue déviants.

## Conclusion

Le succès des méthodes d’évaluation de la beauté des photos est certain. Exposées à un très grand nombre d’images, elles permettent de séparer avec des performances raisonnables, les plus belles des plus médiocres. Nulle doute que ces performances s’amélioreront avec le temps, les travaux qui sont présentés en ce moment ayant encore de nombreuses marges de progrès.

Reconnaissons cependant qu’aujourd’hui leur intérêt réside surtout dans leur capacité à élaborer un premier tri sur de grandes quantités d’images. Si l’on souhaite véritablement distinguer les belles images, il est encore nécessaire de repasser sur ce tri pour sélectionner le petit nombre qui surpasse les autres.

Regrettons, comme nous le faisons pour de nombreux autres problèmes de reconnaissance des formes, que les solutions à base de RNP nous soient livrées sans que soient explicitées les étapes intermédiaires des décisions, ou plutôt que ces résultats intermédiaires, accessibles sous forme de cartes ne soient guère lisibles aujourd’hui avec nos connaissances. Ainsi, si l’on sait à peu près trier les images, on ne sait pas bien vraiment comment ce tri est fait. C’est pour notre entendement un recul par rapport aux approches à base de primitives.

Insistons enfin sur le fait que les méthodes mises en œuvre à ce jour ont totalement ignoré un pan important du jugement esthétique que la littérature met pourtant en avant :

le contexte culturel et socio-éducatif de l'observateur [71]. Cet oubli se comprend car, si l'esthétique est un champ complexe et mal connu, la culture est encore beaucoup plus complexe et très mal modélisée. On ne sait comment la faire intervenir dans nos architectures. Mais cette ignorance nous livre en fait à une culture cachée, celle qui a présidé à l'élaboration des expertises. Dans le cas de la Beauté des images, c'est donc une communauté d'experts ou d'amateurs éclairés en photographie, répartis dans le monde entier, plutôt friands de vie sociale à travers Internet, fervents souvent de technologie, qui sert de référence. C'est là aussi une grande évolution au regard de ce qui a permis de bâtir les socles de l'esthétique. C'est aussi certainement une menace pour des sociétés qui soutiendraient des représentations originales ou marginales.

## Références

- [1] X. Lu, Z. Lin, H. Jin, J. Yang, and J. Wang, "RAPID : Rating pictorial aesthetics using Deep Learning," in *22nd Int. Conf on Multimedia MM'14* (ACM, ed.), (Orlando, Fl.), pp. 457–466, Nov 2014.
- [2] X. Lu, Z. Lin, X. Shen, R. Mech, and J. Z. Wang, "Deep multi-patch aggregation network for image style, aesthetics, and quality estimation," in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 990–998, 2015.
- [3] L. Mai, Y. Niu, and F. Liu, "Composition preserving deep photo aesthetic assessment," in *Proc. IEEE Conf. Comp. Vision and Pattern Recognition*, pp. 497–506, 2016.
- [4] Z. Wang, S. Chang, F. Dolcos, D. Beck, and T. Huang, "Brain-inspired deep network for image aesthetic assessment." arXiv : 1601.04155v2, March 2016.
- [5] Y. Kao, R. He, and K. Huang, "Deep aesthetic quality assessment with semantic information," *IEEE Transactions on Image Processing*, vol. 26, no. 3, pp. 1482–1495, (2017).
- [6] S. Ma, J. Liu, and C. CW, "A-lamp : Adaptive layout-aware multi-patch deep convolutional neural network for photo aesthetic assessment." arXiv : 1704.00248, april 2017.
- [7] H. Talebi and P. Milanfar, "NIMA : Neural image assessment." arXiv : 1709.0541v1, Sept 2017.
- [8] Y. Ke, X. Tang, and F. Jing, "The design of high-level features for photo quality assessment," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, vol. 1, pp. 419–426, 2006.
- [9] R. Datta and J. Wang, "ACQUINE : Aesthetic quality inference engine - real-time automatic rating of photo-aesthetics," in *MIR'10* (ACM, ed.), 2010.
- [10] L. Marchesotti, F. Perronnin, D. Larlus, and G. Czurka, "Assessing the aesthetic quality of photographs using generic image descriptors," *Int. Conf. on Computer Vision*, pp. 1784–1791, 2011.
- [11] N. Murray, L. Marchesotti, and F. Perronnin, "AVA : a large-scale database for for aesthetic visual analysis," in *CVPR : Intern. Conf. on Comp. Vision and Pattern Recognition*, pp. 2408–2415, 2012.
- [12] R. Schifanella, M. Reddi, and L. Aiello, "An image is worth more than thousand favourites : surfacing the hidden beauty of Flickr pictures." ArXiv : 1505.03358v2, May 2015.
- [13] C. Henry, *Introduction à une esthétique scientifique*. Paris : Revue contemporaine, 1885.
- [14] G. Birkhoff, *Aesthetic Measure*. Cambridge (USA) : Harvard University Press, 1933.
- [15] A. A. Moles, "Théorie de l'information et perception esthétique," *Revue Philosophique de la France et de l'Etranger et Flammarion*, pp. 233–242, 1957.
- [16] M. Bense, *Einführung in die informationstheoretische Ästhetik. Grundlegung und Anwendung in der Texttheorie*. Rowoldt Taschenbuch Verlag, 1969.
- [17] J. Rigau, M. Feixas, and M. Sbert, "Information aesthetic measures," *IEEE Computer Graphics and Applications*, vol. 2, pp. 24–34, 2008.
- [18] Platon, *Platon, œuvres complètes*. Flammarion, 2011.
- [19] D. Diderot, *Correspondance littéraire*. Paris : Furne, 1769.
- [20] G. Hegel, *Vorlesungen über die Ästhetik*. 1835-1838.
- [21] E. Kant, *Critique de la faculté de juger*. Librairie J. Vrin (traduction et réédition - 1993), 1790.
- [22] E. H. Gombrich, *Art and Illusion : a study in the psychology of pictorial representations*. Wahington D.C. : Princeton University Press, 1960.
- [23] D. Berlyne, *Aesthetics and Psychobiology*, vol. 336. New-York : Appleton-Century Croft, 1971.
- [24] E. M. Zemach, *La beauté réelle : une défense du réalisme esthétique*. Presses Universitaires de Rennes ; collection "Aesthetica", 2005.
- [25] R. Pouivet, *Le réalisme esthétique*. Paris : Presses universitaires de France, 2006.
- [26] A. Danto, "The artworld," *The journal of philosophy*, vol. 61, no. 19, pp. 571–584, 1964.
- [27] R. Arnheim, *Art and Visual Perception : A Psychology of the Creative Eye*. Los Angeles (CA) : University of California Press, 1954.
- [28] R. Solso, *Cognition and the visual Arts*. Cambridge (Mass.) : The MIT Press, 5th ed., 1996.
- [29] C. Lange and W. James, *The emotions*. Williams & Wilkins, 1922.
- [30] D. Sperber and D. Wilson, "Relevance theory," in *The Handbook of Pragmatics* (G. Horn, L.R. & Ward, ed.), pp. 607–632, Oxford : Blackwell, 2004.
- [31] J. Dessalles, *La pertinence et ses origines cognitives*. Paris : Hermès- Sciences, 2008.
- [32] S. Zeki, *Inner vision : An exploration of art and the brain*. Oxford University Press, 1999.
- [33] A. Chatterjee and O. Vartanian, "Neuroscience of aesthetics," *Annals of the New-York Academy of Sciences*, vol. 1369, pp. 172–194, 2016.
- [34] S. Brown, X. Gao, L. Tisdelle, S. Eickoff, and M. Lotti, "Naturalizing aesthetics : Brain areas for aesthetic appraisal across sensory modalities," *Neuroimages*, vol. 58, pp. 250–258, 2011.
- [35] C. Di Dio and G. Vittorio, "Neuroaesthetics : a review," *Current Opinion in Neurobiology*, vol. 19, pp. 682–687, 2009.
- [36] T. Ishizu and S. Zeki, "Toward a brain-based theory of beauty." PLOS One, July 2011.

- [37] O. Vartanian and M. Skov, "Neural correlates of viewing paintings : evidence from a quantitative meta analysis of fm-ri-maging data," *Brain Cogn.*, vol. 87, pp. 52–56, 2014.
- [38] J. Changeux, *La beauté dans le cerveau*. Paris : Odile Jacob, 2016.
- [39] F. Vidal, "Neuroaesthetics : Getting rid of art and beauty," *Biosocieties*, vol. 7, no. 2, p. 209, 2012.
- [40] S. Brown and E. Dissanayake, "The Arts are more than Aesthetics : Neuroaesthetics as narrow aesthetics," *Neuroaesthetics*, pp. 43–57, 2009.
- [41] P. Bourdieu, *Un art moyen : essai sur les usages sociaux de la photographie*. Paris : Les Editions de Minuit, 1965.
- [42] A. Mittal, R. Soundararajan, and A. C. Bovik, "Making a "complete blind" image quality analyzer," *IEEE Signal Processing Letters*, vol. 20 (3), pp. 209–212, 2013.
- [43] R. Chandler, "Seven challenges in image quality assessment : past, present and future research," *ISRN Signal Processing*, vol. 2013, pp. 1–53, 2013.
- [44] I. McManus, K. Stöver, and D. Kim, "Arnheim's gestalt theory of visual balance : Examining the compositional structure of art photographs and abstract images.," *i-Perception*, vol. 2, no. 6, pp. 615–647, 2011.
- [45] S. Amirshahi, G. Hayn-Leichsenring, J. Denzler, and C. Redies, "Evaluating the rule of thirds in photographs and paintings.," *Art & Perception*, vol. 2, no. 1–2, pp. 163–182, 2014.
- [46] R. Hubner and M. Fillinger, "Comparaison of objective measures for predicting perceptual balance and visual aesthetic preferences," *Frontiers in Psychology*, vol. 7, p. 1 :15, March 2016.
- [47] A. Schweinhart and E. Essock, "Structural content in paintings : Artists overregularize oriented content of paintings relative to the typical natural scene bias," *Perception*, vol. 42, pp. 1311–1332, 2013.
- [48] D. Attewell and R. Baddeley, "The distribution of reflectances within the visual environment," *Vision Research*, vol. 47, no. 4, pp. 548–554, 2007.
- [49] A. Mehrabian, "Individual differences in stimulus screening and arousability," *J. Personality and Social Psychology*, vol. 45, pp. 237–250, 1977.
- [50] D. Smith, "Color-person-environment relationship," *Color Research & Applications*, vol. 33, pp. 312–319, 2008.
- [51] A. Yarbus, *Eye movements and vision*. Plenum Press, 1967.
- [52] R. Rosenberg and C. Klein, *Art, Aesthetics and the brain*, ch. 5 - The moving eye of the beholder : Eye tracking and the perception of paintings, pp. 79–108. Oxford University Press, 2015.
- [53] Y. Luo and X. Tang, "Photo and video quality evaluation : Focusing on the subject," in *ECCV 2008* (D. Forsyth and A. Zisserman, eds.), vol. LNCS 5304, (Berlin-Heidelberg), pp. 386–399, Springer-Verlag, 2008.
- [54] S. Dhar, V. Ordonez, and T. Berg, "High level describable attributes for predicting aesthetics and interestingness," in *Computer Vision and Pattern Recognition (CVPR), 2011* (I. C. on, ed.), pp. 1657–1664, 2011.
- [55] L. Yao, P. Suryanarayan, M. Qiao, J. Z. Wang, and J. Li, "Oscar : On-site composition and aesthetics feedback through exemplars for photographers," *International Journal of Computer Vision*, vol. 96, no. 3, pp. 353–383, 2012.
- [56] K. Lo, K. Liu, and C. Chen, "Assessment of photo aesthetics with efficiency.," in *IEEE International Conference on Pattern Recognition (ICPR)*, pp. 2186–2189., Nov. 2012.
- [57] J. San Pedro, T. Yeh, and N. Oliver, "Leveraging user comments for aesthetic aware image search reranking," in *Proceedings of the 21st international conference on World Wide Web (ACM, ed.)*, pp. 439–448, june 2012.
- [58] P. Lu, Z. Kuang, X. Peng, and R. Li, "Discovering harmony : a hierarchical colour harmony model for aesthetics assessment," in *Asian Conference on Computer Vision*, pp. 452–467, Springer Int., Nov. 2014.
- [59] S. Kong, X. Shen, Z. Lin, R. Mech, and C. Fowlkes, "Photo aesthetics ranking network with attributes and content adaptation.," in *ECCV European Conference on Computer Vision*, pp. 662–679, Springer International Publishing, Oct. 2016.
- [60] Y. Kao, K. Huang, and S. Maybank, "Hierarchical aesthetic quality assessment using deep convolutional neural networks," *Signal Processing : Image Communication*, vol. 47, p. 500 :510, 2016.
- [61] M. Redi, F. Liu, and N. O'Hare, "Bidding the aesthetics gap : The wild beauty of web imaginary," in *ICMR'17 conference (ACM, ed.)*, (Bucarest), pp. 242–250, June 2017.
- [62] R. Datta, D. Joshi, J. Li, and J. Wang, "Studying aesthetics in photographic images using a computational approach," in *Computer Vision, ECCV 2006*, vol. 3953 of *Lecture Notes in Computer Science*, pp. 288–301, 2006.
- [63] K. Lo, K. Liu, and C. Chen, "Intelligent photographing interface with on-device aesthetic quality assessment," in *ACCV 2012 Workshop* (J. Parks and J. Kim, eds.), vol. LNCS 7729, (Berlin), pp. 533–544, Springer Verlag, 2013.
- [64] P. Lu, X. Peng, X. Zhu, and R. Li, "An EL-LDA based general color harmony model for photo aesthetics assessment," *Signal Processing*, vol. 120, pp. 731–745, 2016.
- [65] L. Marchesotti, F. Perronnin, and F. Meylan, "Learning beautiful (and ugly) attributes," in *BMVC*, vol. 7, pp. 1–11, 2013.
- [66] P. Lu, X. Peng, R. Li, and X. Wang, "Towards aesthetics of image : A Bayesian framework for color harmony modeling," *Signal Processing : Image Communication*, vol. 39, pp. 487–498, Nov. 2015.
- [67] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans on PAMI*, vol. 37, no. 9, pp. 1904–1907, 2015.
- [68] R. Diderot, "Le salon de 1967," 1967.
- [69] B. Jin, V. Ortiz-Segovia, and Süssstrunk, "Image aesthetic predictors based on weighted CNNs," in *ICIP, Int. Conf. on Image Processing 2016*, vol. 1, pp. 2291–2295, sept. 2016.
- [70] K. Schwarz, F. Wieschollek, and H. Lensch, "Will people like your image ?," arXiv 1611-05203, University of Tübingen, 2016.
- [71] M. Wascheck, *Le chef d'œuvre : un fait culturel*, ch. 2, pp. 25–46. Qu'est-ce qu'un Chef d'œuvre ?, Gallimard (Paris), 2000.