



Analyser l'archive visuelle par l'image

Les méthodes qualitative et statistique en contraste

Organisé par :

Maria Giulia Dondero (FNRS/ULg)

Maud Hagelstein (FNRS/ULg)

Thierry Lenain (ULB)

Jean Cristtus Portela (UNESP-Araraquara/CNPq)

8 et 9 décembre 2016

Salle Lumière/salle de l'Horloge

Université de Liège



Argumentaire

La sémiotique visuelle a privilégié majoritairement l'analyse d'images singulières et exemplaires, en développant une approche herméneutique (ou « rapprochée »). Là où l'histoire de l'art s'est pour une part consacrée à l'évolution des styles (cf. les travaux de A. Riegl, H. Wölfflin, A. Warburg ou H. Focillon), la sémiotique visuelle a toujours favorisé l'analyse synchronique.

Dans le domaine des sciences du langage, des méthodes alternatives (notamment quantitatives) permettent d'étudier des mouvements et des corpus plus vastes, en visant l'exhaustivité plutôt que l'exemplarité et l'analyse diachronique. Si la fortune de l'analyse automatique de corpus verbaux par la linguistique est bien connue, les analyses d'ensembles d'images plus importants n'ont pas encore fait l'objet de débats substantiels au sein des sciences humaines et des sciences du langage dans l'Europe francophone.

L'objectif principal de ce colloque sera de tester la proposition d'analyse statistique de corpus d'images dans la diachronie formulée par Lev Manovich. L'approche de la « Media Visualization » de Manovich favorise la production et la théorisation de descriptions visuelles de larges corpus d'images provenant de médias divers (peinture, photographie, BD, jeux vidéo, films), archivées en ligne ou stockées dans des bases de données spécialisées. Si les analyses de corpus visuels, contrairement à celles de corpus verbaux, ne permettent pas d'isoler des unités discrètes, Manovich teste néanmoins des paramètres suprasegmentaux pour étudier les collections d'images, les rendre commensurables et enfin transposables en visualisation globale.

Les images sont ainsi filtrées selon des paramètres visuels quantifiables — la teinte, la saturation, la luminosité, les niveaux de gris, la position, la dimension de formes, etc.

Cette rencontre, qui rassemblera linguistes, sémioticiens, historiens et philosophes de l'art, ainsi que des scientifiques de la communication, permettra de comparer les différentes méthodes d'analyse des images : en particulier, il s'agira de confronter les approches qualitatives et quantitatives pour essayer de comprendre leurs croisements, leurs conflits, leurs complémentarités.

Comité scientifique

Sémir Badir (FNRS/ULg)
Pierluigi Basso Fossali (Université de Lyon II-Lumière)
Anne Beyaert-Geslin (Université Bordeaux Montaigne)
Marion Colas-Blaise (Université du Luxembourg)
Maria Giulia Dondero (FNRS/ULg)
Maud Hagelstein (FNRS/ULg)
Thierry Lenain (ULB)
Odile Le Guern (Université Lyon II-Lumière)
Jean-Marie Klinkenberg (ULg)
Jean Cristtus Portela (UNESP Araraquara/CNPq)
François Provenzano (ULg)
Everardo Reyes (Université Paris 8)

Programme

8 décembre 2016, Salle Lumière

9h00-9h30 Allocution du Directeur de l'U.R. Traverses, Jean-Pierre Bertrand (ULg)

9h30-10h00 Maria Giulia Dondero (FNRS/ULg)
Introduction

10h00-10h45 François Rastier (CNRS/INALCO)
Genres d'images et sémiotique de corpus

Pause

11h00-11h45 Everardo Reyes (Université Paris 8)
Enjeux des cultural analytics et sémiotique visuelle

11h45-12h30 Clarisse Bardirot (Université de Valenciennes)
Big data is beautiful/ big data is meaningful

Déjeuner

14h30-15h15 Peter Stockinger (INALCO Paris)
Pour une approche sémiotique des archives audiovisuelles

15h15-16h00 Alessandro Zinna (Université de Toulouse 2 - Jean Jaurès)
Le corpus Francis Bacon : une archéologie

Pause

16h20-17h05 Dario Compagno (Université Paris 3 – Sorbonne Nouvelle) et Céline Cholet (Université Bordeaux Montaigne)
Valider l'immanence. Quelles pistes possibles pour une mise à l'épreuve empirique des hypothèses sémiotiques sur le sens les images ?

17h05-18h00 Claudio Paolucci (Université de Bologne)
Archive, patrimoine et mémoire. La tiers-mondialisation du savoir à l'ère de la numérisation

9 décembre, Salle de l'Horloge

9h30-10h15 Marion Colas-Blaise et Emmanuelle Pelard (Université du Luxembourg)
Quand le qualitatif rencontre le quantitatif : visualisation de l'information, densité et diagrammaticité

10h15-11h00 Pierluigi Basso Fossali (Université Lyon II-Lumière)
Rencontrer une condition iconique. Archiver la densité, reconstruire les approches des images

Pause

11h30-12h15 Enzo D'Armenio (Université de Bologne)
L'analyse statistique de l'archive à la lumière du montage intermédial : les formats techniques des images

Déjeuner

14h15-15h00 Jeremy Hamers (ULg)
Notes équivoques pour une critique marxiste de l'analyse statistique des images (H. Farocki / H.M. Enzensberger)

15h00-15h45 Jean Cristtus Portela (UNESP-Araraquara/CNPq)
Analyser (un corpus de) la bande dessinée à l'âge du numérique

16h00 Clôture

Analyser l'archive visuelle par l'image : les méthodes qualitative et statistique en contraste

Université de Liège, les 8 et 9 décembre 2016, salle Lumière/salle de l'Horloge

Résumés des interventions

Clarisse Bardiot (Université de Valenciennes)

Big data is beautiful/ big data is meaningful

Le programme *Culture analytics* a réuni au printemps 2016 à UCLA des chercheurs en sciences humaines, des mathématiciens et des informaticiens. La définition à laquelle nous avons aboutie, « culture analytics is the data-driven analysis of culture » considère la culture au sens large du terme : il s'agit autant des interactions sociales que des expressions culturelles. La visualisation des données représente à la fois un enjeu fondamental pour les *Culture analytics*, qu'il s'agisse de corpus de textes, d'images, ou encore de relations sociales, et un écueil : parfois, la séduction des images produites l'emporte sur leur signification. De ce point de vue, la démarche de Lev Manovich est exemplaire : les visualisations sur lesquelles il appuie ses analyses de larges corpus d'images sont également exposées dans des galeries et présentées comme des œuvres. Par ailleurs, alors que les premiers résultats produisaient des images fixes dans lesquelles il faut zoomer pour avoir accès à la vue précise d'éléments du corpus, le spectateur est de plus en plus souvent invité à s'emparer de visualisations interactives dont il doit déduire le sens. Je reviendrai sur ces deux phénomènes à partir d'exemples tirés des travaux de Lev Manovich et d'autres chercheurs travaillant sur l'analyse de corpus visuels.

Pierluigi Basso Fossali (Université Lyon II-Lumière)

Rencontrer une condition iconique. Archiver la densité, reconstruire les approches des images

L'image est toujours prise entre ses figements textuels et ses dérives imaginatives. Sa déclinaison sensible immédiate convainc le regard de sa maîtrise totale d'une région du visible, mais cette juridiction apparente n'est qu'un leurre qui finit par transformer le soupçon d'une densité iconique résistante en une prolifération factice d'autres pertinences possibles.

Les nouveaux médias n'ont fait qu'accentuer cette ambiguïté entre une image-écran disponible au monitoring et une image-écran susceptible d'ouvrir d'autres trames indiciaires, mais encore insondables.

Archiver la densité de l'image est alors un paradoxe : cette demande est à la fois *inutile*, vu que l'image contient des « photographies » pas encore développées du sens, et *désespérée*, car aucune organisation de métadonnées pourra jamais établir non seulement un plan unique de pertinences thématiques, mais non plus la généalogie diagrammatique à laquelle on voudrait la ramener.

La prolifération de visualisations d'une image en tant que paquet de données n'empêche pas qu'il y ait encore des arrêts sur des icônes électives,

lesquelles invitent le « corps social » à reconnaître ce qui résiste au flux des consommations immédiates.

De manière plus générale, si progressivement on a porté l'attention sur la relation entre les images et les pratiques qui en ont motivé la production, il faudra préciser aussi le régime de jouissance caractérisant les corpus *emic* (appréhension synoptique, compulsation séquentielle, implémentation en affiche, etc.), aussi que les critères pour l'établissement des corpus *etic*, sous l'impulsion de l'iconologie, des études de tendances, de l'archéologie représentationnelle, etc. On peut se poser la question enfin si le numérique nous aide à reconstruire de cadres d'interprétation historiquement adoptés (les « potentialisations passées ») ou s'il nous conduit en revanche à virtualiser à nouveau des paradigmes d'approches du visible qui ont toutes les possibilités ouvertes, sauf celle de préserver le mobile recélé d'une image émergente.

Marion Colas-Blaise et Emmanuelle Pelard (Université du Luxembourg)
Quand le qualitatif rencontre le quantitatif : visualisation de l'information, densité et diagrammaticité

Cette communication se proposera de réfléchir à la manière dont notre rapport à la peinture (en particulier celle de Piet Mondrian) est reconfiguré par les techniques de visualisation de l'information proposées surtout par Lev Manovich. Si, d'une part, la visualisation de données complexes suppose la discrétisation et l'extraction de propriétés du tableau telles que la teinte, la luminosité, la saturation et compromet ainsi une approche herméneutique, et si, d'autre part, elle semble privilégier le langage de l'abstraction (à travers la transposition en langage algorithmique), il faudra se demander en quoi elle permet de conférer aux tableaux un surcroît de sens.

À cet effet, nous mettrons en avant l'introduction de la variable temps, d'un côté, et la création d'une plus grande densité et d'une plus grande complexité, de l'autre. Il s'agira, ensuite, de revoir, à la lumière de ces considérations, voire de dépasser la distinction goodmanienne entre les œuvres autographiques et allographiques à travers, essentiellement, la notion de (contexte d') implémentation.

Globalement, nous chercherons à vérifier l'hypothèse de Manovich selon laquelle la visualisation de l'information relève conjointement de la science, du design et de l'art.

Enzo D'Armenio (Université de Bologne)

L'analyse statistique de l'archive à la lumière du montage intermédial : les formats techniques des images

Notre proposition explore la convergence possible entre l'approche qualitative de la sémiotique et les techniques statistiques pour l'analyse des archives. En examinant une série de cas audiovisuels qui associent des substances de l'expression différentes (Dondero et Reyes-Garcia 2016), nous proposons un concept potentiellement transposable à l'analyse informatisée : le format technique des images. Par cette notion, nous désignons les composants de l'expression qui renvoient au dispositif et aux coordonnées spatio-temporelles de la production : la définition, l'espace de représentation et la fluidité font partie des paramètres qui les caractérisent. Nous tenterons de démontrer que ces variables peuvent non seulement jouer un rôle utile dans la théorie de l'énonciation en sémiotique (Benveniste 1966-1974 ; Greimas et Courtés 1979 ; Metz 1991), mais aussi être quantifiées pour une utilisation dans les analyses informatisées (Douglas, Huber et Manovich 2011).

La première partie de l'intervention propose une définition du concept de format technique, en examinant une série d'exemples cinématographiques qui utilisent des formes particulières de montage intermédial (Montani 2010). En mélangeant dessin animé et image photographique, haute et basse définition, reconstruction fictionnelle et matériaux d'archive, ces cas nous permettront d'évaluer la fonction critique des formats techniques et d'en généraliser les propriétés. Dans la deuxième partie nous proposons une exploration des techniques d'analyse quantitatives à partir des recherches de Lev Manovich (2015). Le concept de *semantic gap* nous permettra de réfléchir sur le rôle possible de la sémiotique dans le paradigme de la *Data Science* : cette notion vise à aborder la différence cognitive qui sépare l'interprétation humaine et le fonctionnement des applications informatisées. Les humains reconnaissent les visages et comprennent des textes verbaux, tandis que les applications d'analyse voient quantitativement les différences entre groupes verbaux et formations visuelles. La possibilité informatique d'élaborer des grandes quantités de données retrouve une limite sémiotique, parce que la capacité d'associer des éléments expressifs et des éléments du contenu est encore limitée (Redi et al. 2016).

La sémiotique peut toutefois élaborer des propositions en liant la quantification du plan de l'expression aux éléments de l'énonciation productive : un examen des paramètres majoritairement utilisés dans l'analyse quantitative des images – paramètres comme la saturation, la luminosité, la quantité et la complexité des formes, la longueur des cadrages, leur mouvement (Manovich 2013) – rend concrète la possibilité d'élaborer des algorithmes pour l'individuation des formats techniques et nous permet d'en prévoir l'efficace.

Bibliographie

- Benveniste, É. (1966-1974) *Problèmes de linguistique générale*, 1-2, Paris: Gallimard.
- Dondero, M.G. et Reyes-Garcia, E. (2016), « Les supports des images : de la photographie à l'image numérique », *Revue française des sciences de l'information et de la communication* 9, [En ligne]. URL : <https://rfsic.revues.org/2124>
- Douglas, J., Huber, W., Manovich, L. (2011), « Understanding scanlation: how to read one million fan-translated manga pages », *Image & Narrative*, Vol 12, No1.
- Greimas, A.J. et Courtés, J. (1979) *Sémiotique. Dictionnaire raisonné de la théorie du langage*, Paris: Hachette.
- Manovich, L. (2013), « Visualizing Vertov », *Russian Journal Of Communication* Vol. 5, 1.
- Manovich, L. (2015), « Data Science and Digital Art History », *DAH-Journal*, 1.
- Metz, C. (1991), *L'énonciation impersonnelle, ou, le site du film*, Paris : Méridiens Klincksieck.
- Montani, P. (2010), *L'immaginazione intermediale. Perlustrare, rifigurare, testimoniare il mondo visibile*, Roma-Bari: Laterza.
- Redi, M., Crockett, D., Manovich, L. et Osindero, S. (2016), « What Makes Photo Cultures Different », *MM'16 Proceedings of the 2016 ACM on Multimedia Conference* : 287-291.

Claudio Paolucci (Université de Bologne)

Archive, patrimoine, mémoire. La tiers-mondialisation du savoir à l'ère numérique

Ce travail vise à montrer les rapports entre les notions d'archive, patrimoine et mémoire en mettant en avant ce qui a changé dans l'idée foucauldienne d'archive à l'ère de la numérisation et du World Wide Web. En se concentrant sur la loi de la rareté des énoncés, qui semble avoir été invalidée par la numérisation, on propose ici une reformulation sémiotique du concept d'archive, en utilisant les notions sémiotiques d'Encyclopédie et de Semiosphère.

Jean Cristtus Portela (Unesp-Araraquara/CNPq)

Analyser (un corpus de) la bande dessinée à l'âge du numérique

Comment peut-on constituer un corpus de bande dessinée ? Cette question un peu naïve, qui ne laisse pas pourtant les sémioticiens indifférents, devient un choix de principe dès que nous dépassons les limites de l'analyse d'une planche et d'un album, qui s'accommodent le plus souvent de visées qualitatives et synchroniques, et nous abordons des centaines, voire des milliers de planches sollicitant des approches quantitatives, diachroniques et comparatives visant à une histoire des formes. L'analyse d'un corpus restreint de BD se contente le plus souvent de se borner par des principes discursifs orientés par le plan du contenu tels que la thématique, la narrativité, la figurativité et les modes d'énonciation. Tout se passe autrement si nous abordons un corpus comprenant une masse importante de planches à

analyser, dans la mesure où l'approche qualitative de la sémiotique de la bande dessinée, basée sur des saisies intelligibles des ensembles signifiants, se révèle stérile face à la profusion de sémiotiques-objets à viser et à analyser. Dans les domaines élargis de gros corpus de BD (par exemple, les strips américains du début du XX^{ème} siècle, la BD belge des années 1950 ou la BD brésilienne contemporaine, etc.), qui nous obligent à gérer la complexité d'une diversité de manifestations textuelles objectivables, l'approche du contenu doit être mise entre parenthèses à la faveur d'une approche de l'expression qui prend en compte des paramètres sensibles relevant de la forme, de la couleur, de la topographie, de la focalisation, etc., des BD qui « font corpus ».

Notre réflexion sur la façon de construire des corpus élargis de BD portera essentiellement sur deux types d'outils numériques de visualisation : 1) Des moteurs de recherche en ligne (*Google Images* ou le site spécialisé *Bédéthèque*) qui opèrent à travers des repères lexicaux de textes verbaux (titres, liens, menus, légendes et texte explicatifs, entre autres) et accompagnent les images et des métadonnées préétablies (titre, crédits de scénario, dessin et couleurs, dépôt légal, impression, prix moyen, éditeur, format, numéro de planches, résumé, etc.) ; 2) Le projet « One million manga pages », dirigé par Lev Manovich au sein du laboratoire *Software Studies Initiative*, qui propose l'organisation et l'analyse d'un corpus de 883 séries de manga comprenant plus d'un million de pages, extraites du site web *OneManga.com*, réunissant des mangas numérisés, traduits et distribués par des fans (la « scanlation »). Notre but majeur est de comprendre comment ces différentes manières de rassembler et de constituer des corpus d'images de bandes dessinées impliquent des stratégies discursives et textuelles de délimitation, de segmentation et de classement qui peuvent être mises au service de l'analyse sémiotique, et changer notre manière d'analyser la bande dessinée.

François Rastier (CNRS, INALCO)

Genres d'images et sémiotique de corpus

À l'heure actuelle, dans diverses disciplines, de la linguistique à la musicologie, de l'iconologie aux études de cinéma, de l'archéologie à l'anthropologie, des corpus d'étendue et de qualité croissante sont constitués et leur investigation est assistée par divers logiciels.

Or : (i) La dimension appliquée et l'instrumentation raisonnée exigent des clarifications méthodologiques. (ii) Le choix du corpus impose une perspective critique. (iii) Le choix des tests instrumentés redouble les problèmes herméneutiques. (iv) L'annotation (métadonnées et balisage) renoue avec la problématique philologique.

Ces réflexions seront appliquées à la *photo people* dans les magazines et aux photos touristiques dans les catalogues d'agences de voyage. Dans ces

deux domaines, l'objectif poursuivi par l'analyse des genres d'images est l'extraction de traits génériques caractérisants pour l'aide à l'indexation et l'indexation automatique.

Everardo Reyes-Garcia (Université Paris 8)

Enjeux des cultural analytics et sémiotique visuelle

Nous concevons les apports du courant nommé "Cultural analytics" à l'instar d'une ressource méthodologique pouvant compléter les méthodes en sémiotique picturale et des images. Dans ce sens, nous présentons des exemples de modèles de représentation de collections d'images. Au travers ces explorations nous soulevons des critiques et des problèmes quant à l'adoption de l'approche quantitative de la visualisation des données visuelles, plus particulièrement la description des formes, la réorganisation des totalités, l'interactivité et navigation.

Alessandro Zinna (Université de Toulouse 2 – Jean Jaurès)

Le corpus Francis Bacon : une archéologie

La publication récente du catalogue raisonné des œuvres de F. Bacon, publié par Martin Harrison en 2016, nous donne une vision globale de sa production. Ce corpus, qui se veut exhaustif, se prête bien à une analyse quantitative et/ou qualitative. Il pourrait être soumis au traitement Manovich pour obtenir des données statistiques par rapport à des paramètres de teinte, saturation, luminosité, couleurs, position des figures, dimensions. Pourtant c'est la reconnaissance des *formes* qui pose problème dans le traitement automatique des images de F. Bacon. Si le problème de la *variation linguistique* est résolu par la lemmatisation, les images posent la question de comment reconnaître la récurrence d'une forme quand elle est *peinte en variation*. Bien que la séduction d'un traitement automatique soit forte, le corpus de peintures de F. Bacon demande une analyse qualitative pour découvrir la récurrence des figures peintes selon le régime de la *déformation*. Plusieurs éléments nous indiquent que le corpus Bacon est un corpus organique qui fait évoluer d'un tableau à l'autre la représentation des figures. Paradoxalement, une analyse quantitative manquerait la récurrence la plus importante de son œuvre. Il serait pourtant une erreur de considérer ces tableaux dans la *diachronie* car il n'y a pas de diachronie dans la mise en variation d'une figure : la seule voie est celle de retracer le lien de dérivation par la méthode de l'*archéologie*.