



Mémoire de Master 2 MIAGE

COMMENT, À PARTIR DE LA RECONNAISSANCE FACIALE
ET DE LA FOUILLE DE DONNÉES, PEUT ON IDENTIFIER
L'ENSEMBLE DES SOSIES POTENTIELS D'UNE PERSONNE ?

Application à l'univers du cosplay

Réalisé par
VILLER Nathanaëlle
Tuteur
GUEHIS Sonia

Année scolaire 2017-2018

Table des matières

1	Description du contexte	1
2	Liaison avec l'informatique	2
3	Logiciels existants	3

1 Description du contexte

Le phénomène du cosplay a commencé dans les années 40 aux Etats Unis. Mais c'est lors des années 70 et 80 que le cosplay atteint son paroxysme grâce aux succès de Star Trek et Star Wars. Le Japon va ensuite adopter ce phénomène et lui donner le nom qu'il porte aujourd'hui : le cosplay (qui vient de la contraction de costume et playing). Le principe est simple : Il faut se déguiser en un personnage de fiction et incarner ce personnage. Le mouvement va s'intensifier au Japon et Tokyo deviendra le centre névralgique de ce phénomène. Il faudra attendre les années 90 pour que le cosplay arrive en Europe. La scène européenne se démarquera par sa volonté de respect du craft (fabrication manuel de son costume). Les conventions et concours vont se multiplier. Les cosplayers vont de plus en plus chercher de nouvelles techniques pour pouvoir ressembler le plus possible aux personnages qu'ils essaient d'incarner.

Certains cosplayers se démarqueront et deviendront professionnels dans le milieu mais pour monsieur tout le monde c'est plus difficile. Il y a parmi eux 2 grandes catégories qui se démarquent. Ceux qui cosplay leurs personnages préférés sans se soucier de la morphologie et ceux qui essaie de trouver des personnage qui leur correspond physiquement. Certains cosplayer essaie de trouver de nouveaux personnage en faisant des appels sur leur page "qui pourrais je incarner ? " ou en demandant dans des groupes "connaissez vous un personnage blond ... ". Mais pour trouver de nouvelles inspirations qui leur correspondent ce n'est pas facile.

La recherche est longue et parfois un peu décourageante pour certains. Lorsque l'objectif est de pouvoir ressembler au maximum à un personnage l'idéal serait de trouver un personnage qui nous ressemble un maximum au niveau de la morphologie.

De plus les cosplayers cosplayent des personnages qui leur plaisent, leur parlent ou qui sont d'un univers qu'ils aiment. De nos jours nous avons accès à tout un panel d'informations par internet, et à mesure que le phénomène grandi il est difficile de faire le tri , de rechercher des données et d'avoir la capacité de traiter l'ensemble des informations récoltées.

De manière à améliorer cette recherche, il serait possible d'adapter de nouveaux usages florissant ces dernières années de manière à optimiser et automatiser ce type de fouille.

Comment l'utilisation des outils et pratiques récentes en informatiques permettraient de faciliter la recherche de personnages correspondant morphologiquement a une personne réelle ? Et, dans un deuxième temps, d'apprendre de ses goût afin de proposer à l'utilisateur un contenu personnalisé ?

2 Liaison avec l'informatique

Récemment j'ai découvert le crawling web, ou en français, le robot d'indexation. Il s'agit de bots logiciels qui vont parcourir les sites internet avec pour objectif principal de les indexer. Le parcours se fait en profondeur, le logiciel accédant tour à tour à chaque pages ou sites externes grâce aux liens hypertextes présents dans la page en cours d'analyse (exemple : lien url) puis réitère l'opération. On peut ajouter des paramètres au bot afin qu'il recherche en même temps que sa "lecture". Il m'est donc venu à l'idée d'appliquer le concept du robot d'indexation pour aider à la recherche massive des univers ou personnages fictifs afin de proposer aux cosplayers les informations trouvées et les aider à trouver de nouveaux projets. Ainsi en fonction de paramètre clair et définis le bot parcourra la toile à la recherche d'un personnage idéal.

Mais pourquoi simplement s'arrêter à la chercher de personnages en fonction de certains critères d'appréciation tels qu'une couleur de cheveux, ou un genre ? En poussant un peu plus loin ma réflexion il met venu à l'idée d'incorporer à mon champs de recherche la reconnaissance faciale qui permettrait d'analyser une, ou même plusieurs, photo de l'individu et ainsi décrypter ses critères morphologiques. Couleur de peau, mensurations, âge, forme du visage, etc ... Et élargir ainsi les éléments de recherche du bot afin de chercher non seulement un personnage qui pourrait nous plaire mais aussi un personnage qui nous ressemble physiquement.

L'outil ainsi présenter pourrait encore s'améliorer, nous pourrions imaginer que l'outil de lui même, par habitude, puisse libérer l'utilisateur de la contrainte d'informer de ses changements physiques ou à modifier les secteurs de recherche liés à ces centres d'intérêts. Grâce au machine learning il devraient être possible de suivre l'évolution de l'utilisateur. Grâce à sa page facebook on pourrait récupérer les cosplays déjà réalisés et ainsi connaître les univers qui lui plaisent. Et grâce à des sites culturels comme sens critique qui permettent de lister les films vus, les livres lus, les musiques écoutés, de les noter et même de mettre ce qu'on aimerait voir, lire ou entendre, le bot pourrait apprendre de l'utilisateur et ainsi affiner les paramètres de recherches. L'utilisateur n'aura ainsi plus besoin de venir mettre à jour ses informations car le bot apprendra de lui même à force d'utilisation.

Ce type d'outil serait un énorme avantage, il est très contraignant pour un utilisateur de devoir mettre à jour soi même les informations. D'autant plus que l'utilisateur a tendance à oublier des choses, ou ne pas avoir une idée précise de ce qu'il souhaite trouver.

3 Logiciels existants

On trouve facilement plusieurs crawler tel que :

- **GoogleBot** est le crawler le plus connu. Il permet d'indexer les sites internet. Il fait principalement de la "lecture" de site. Ce qui, dans notre cas est insuffisant car l'objectif est de "chercher" et non seulement de lire les sites.

Dans le même ordre d'idée nous pouvons retrouver les bots des moteurs de recherche tel que BingBot, Slurp Bot (Yahoo), DuckDuckBot, BaiduSpider (Baidu est un moteur de recherche en chinois) ou encore YandexBot (utilisé pour les moteurs de recherches russes).

- **Facebook External Hit**, Facebook utilise ce crawler afin de permettre à ces utilisateurs, lorsqu'ils envoient (publient) un lien internet à un autre utilisateur du site, de récupérer et afficher certaines informations importantes comme par exemple : des extraits vidéos, des images ou des titres.

Ce bot ne correspond pas à ce que nous souhaitons mettre en place. Par contre il pourrait être un élément de réponse à la question des droits d'auteurs. En effet, si l'outil était développé concrètement comment renvoyer l'image trouvé ainsi que toutes ses informations sans enfreindre les lois sur le droit à l'image ? Avec ce type de bot cela est possible car il permet de récupérer et afficher l'information choisi et de renvoyer vers sa localisation dans la toile.

- **HTTrack** est un logiciel aspirateur de site internet qui crée des miroirs des sites Web pour une utilisation hors ligne.
- **Alexa Crawler** (développé par Amazon) va indexer les sites après les avoir triés en fonctions de certains critères.

Cela se rapproche de ce que nous souhaitons en apportant une dimension de classement selon des critères multiples qu'il est possible de sélectionner et filtrer.

- **Nutch** est un crawler destiné à un moteur de recherche open source. Son architecture est modulaire et permet de créer des plugins pour différentes phases du processus : récupération des données, analyse des documents, recherche, etc.

Nutch est le crawler qui se rapproche le plus de la solution que nous cherchons.

Pour ce qui est des plugins de machine learning il existe Mahout qui est une bibliothèque open source de Machine Learning plutôt complète, très connu et recommandé. Ses algorithmes sont bien testés et supportés. Elle correspond sans soucis à nos besoins.

Concernant la reconnaissance faciale et morphologique il existe à l'heure actuelle de nombreux logiciels permettant de réaliser ce type d'analyse pour les visages mais peu d'information et d'outils sont accessibles pour la reconnaissance morphologique. Nous allons avoir besoin de chercher un logiciel permettant d'analyser une image et d'y détecter un visage réel comme imaginaire et d'en déduire sa morphologie.