

Présentation de l'article *Learning to tag*
(Lei Wuk, Linjun Yang, Nenghai Yu,
Xian-Sheng Hua, 2009)
UE APAS M2 IAD

Jérémie DECOCK Daniel RISTIC

Master d'Informatique
Université Pierre et Marie Curie

21 octobre 2009

Plan

- 1 Le problème
- 2 La solution proposée
- 3 Les résultats expérimentaux
- 4 Critiques

Plan

- 1 Le problème
- 2 La solution proposée
- 3 Les résultats expérimentaux
- 4 Critiques

Pourquoi étiqueter ?

On souhaite pouvoir faire une recherche par mots clés sur des images.

- Une façon d'y arriver : ajouter des tags
- Ne peut pas se faire automatiquement

Sur des sites comme *Flickr*

- Énormément d'images (9 milliards en 2009)
- Fait par les utilisateurs (social tagging)

Étiqueter correctement est difficile

Intuitivement, un bon étiquetage est :

- **complet** tous les éléments de l'image sont documentés par au moins un tag.
- **précis** les termes utilisés sont les termes les plus spécifiques pour l'objet décrit.

But : améliorer la qualité des tags dans la collection d'images

Une des façons de faire est d'aider l'utilisateur en le guidant au moment où il saisit les tags.

- recommandation interactive
- facilite le choix de bons tags

Plan

- 1 Le problème
- 2 La solution proposée**
- 3 Les résultats expérimentaux
- 4 Critiques

La solution

Elle consiste en deux étapes :

- 1 Appliquer trois modèles indépendants
- 2 Combiner ces modèles pour obtenir la "bonne" recommandation

Les modèles

Tag co-occurrence (TC)

Mesure à quel point deux différents tags sont corrélés

$$R_{tag}^a(t_i, t_j) = \frac{|t_i \cap t_j|}{|t_i|}$$

$$R_{tag}^s(t_i, t_j) = \frac{|t_i \cap t_j|}{|t_i \cup t_j|}$$

Tag content correlation (TCC)

Mesure la corrélation des tags en regardant à quel point les images qu'ils décrivent sont similaires entre elles

Image conditioned tag correlation measurement (ITC)

Mesure la corrélation des tags en regardant à quel point les images qu'ils décrivent sont similaires à l'image cible

La combinaison

Algorithme Rankboost (source : wikipedia)

Valeurs d'entrée

Soit un ensemble d'apprentissage annoté : $(x_1, y_1), \dots, (x_m, y_m)$ où $x_i \in X$, sont les exemples et $y_i \in Y = \{-1, +1\}$ les annotations. On notera i_p l'indice des exemples positifs et i_n ceux des exemples négatifs.

Initialisation

On initialise la distribution des exemples par

$D_1(i_p, i_n) = \frac{1}{np*nn}$, $i = 1, \dots, m$. avec np le nombre de positif et nn le nombre de négatif

La combinaison

Algorithme Rankboost (suite)

Déroulement

Pour $t = 1, \dots, T$:

- Trouver le classifieur h_t qui maximise le score de classification en fonction de la difficulté des exemples : D_t

$$r_t = \arg \max_{h_t \in \mathcal{H}} \sum_{i_p, i_n}^m D_t(x_{i_p}, x_{i_n}) [h_t(x_{i_p}) - h_t(x_{i_n})]$$
- On choisie alors le poids du classifieur : $\alpha_t \in \mathbf{R}$, avec

$$\alpha_t = \frac{1}{2} \ln \frac{1+r_t}{1-r_t}$$
- On met ensuite à jour la pondération des couples d'exemples d'apprentissage

$$D_{t+1}(x_{i_p}, x_{i_n}) = \frac{D_t(x_{i_p}, x_{i_n}) e^{-\alpha_t(h_t(x_{i_n}) - h_t(x_{i_p}))}}{Z_t}$$

avec Z_t un facteur de normalisation

La combinaison

Algorithme Rankboost (fin)

Résultat

Le classifieur résultant du processus de sélection est :

$$H(x) = \sum_{t=1}^T \alpha_t h_t(x)$$

Plan

- 1 Le problème
- 2 La solution proposée
- 3 Les résultats expérimentaux**
- 4 Critiques

Ce qu'on veut tester

Les tests doivent notamment vérifier

- que la méthode est plus efficace que ce qui existe actuellement
- que la méthode est robuste
- que le temps d'exécution permet l'interactivité

Le protocole de tests

- 1 million d'images extraites de *Flickr*
- 20 000 tags associés
- on applique une méthode
- on a une liste ordonnée de tags (par ordre d'importance)
- pour les N premiers, on demande à un opérateur humain de répondre à la question "le tag est il pertinent pour l'image considérée ?"

De toutes ces réponses binaires on tire des mesures pour la précision et la couverture

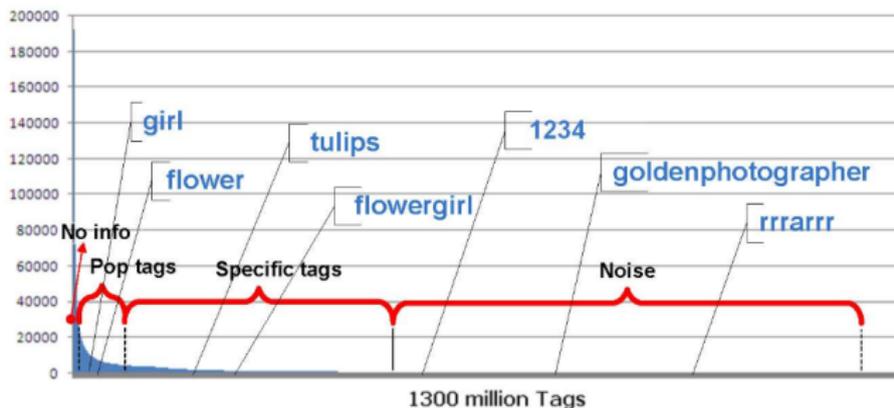
Les résultats de tests

- TC vs (TC+TCC+ITC+Rankboost)
- combinaison linéaire vs rankboost
- variation du nombre initial de tags
- robustesse (50% de tags non pertinents)
- temps d'exécution (< 1 seconde)

Plan

- 1 Le problème
- 2 La solution proposée
- 3 Les résultats expérimentaux
- 4 Critiques**

Quelques approximations



Seuls les tags apparaissant entre 50 et 20000 fois dans la base des images étudiées sont pris en compte.

Le choix de ces deux bornes semble totalement arbitraire car les auteurs ne le justifient à aucun moment.

Critique du protocole de tests

- Les protocoles de tests s'appuient sur un étiquetage humain pour évaluer la couverture des différents modèles. La pertinence d'un tag associé à une image est subjective. Rien n'est dit quand à la redondance de ces évaluations
- Le protocole de test décrit la façon de calculer la couverture mais ne dit pas à quoi correspond la précision ni comment elle est calculée
- Les auteurs présentent les modèles symétriques et asymétriques mais au final, ils n'indiquent pas clairement quel modèle est utilisé dans les expériences

Problèmes non soulevés



Ajoutée le 12 novembre 2007
par [Paolo Margari](#) | [paolomargari.it](#)

Tags

- gatto
- cat
- micio
- kit
- kitty
- pet
- pets
- animali
- animals
- animal
- funny
- hidden
- nascosto
- scared
- impaunito
- cats
- chat
- chats
- nature
- life

Informations supplémentaires

Certains droits réservés (licence
Creative Commons)

- Une image peut être décrite par des mots de différentes langues
- Le mélange des langues risque de poser un problème avec le modèle proposé par les auteurs et ce problème n'est traité à aucun moment

Fin de la présentation

Vos questions !