

AUDIBLEPHOTO

2008年度下期 未踏IT人材発掘・育成事業

Enhance Digital Imaging Experience With Embedded Ambient Sound

クリエイター : アディヤン ムジビヤ
PM : David J Farber

ABSTRACT
本プロジェクトの目的は、拡張されたExchangeable Image File Format (EXIF)、もしくはID3を用いることで、環境音声が付付けられたデジタル写真を構成することである。本プロジェクトはスマートフォン (iPhone 3G/3GS) を基本プラットフォームとして開発を進めた。結果として新しい感覚のデジタル・フォト・エクスペリエンスを可能にし、また、デジタル写真の形式を劇的に再定義することで、この分野において新しいフィールドを開拓することが期待出来る。
本プロジェクトの一環として、オーディオ風景をベースにした写真の自動タギングを提案した。オーディオ風景のタグを用いた写真のグルーピングやインデクシングを開発した。Picasa やFlickr といった写真共有サービスは既にEXIFタグをサポートする。これは、本プロジェクトにとって強力な基盤プラットフォームになることは間違いないであろう。更なる展開として、GPS タグやコンパス・データ等をEXIF タグに埋め込むことによってデジタル写真の新しいビジュアライゼーションの可能性を検討した。

基本アーキテクチャ

- 周囲の音声を写真に付加した
- カメラのシャッターが切れる直前と直後のタイミングで周囲の音声を録音
- 写真と音声は一体化したファイルに保存



他のデータを写真に付加する

- 環境音声の他に、GPS, WiFi Fingerprint, デジタル・コンパスをEXIFデータに追加する
- 他の技術 (PlaceEngine等) への連携が期待出来る
- 写真のビジュアライゼーションを支援する

写真+音声ファイル形式

EXIFを用いたアプローチ

- 他のEXIFタグ (GPSの緯度経度、タグ、ユーザコメント等) をそのまま利用出来る
- 写真がメインのデータになっており、解像度の制限無し

ID3を用いたアプローチ

- 写真観覧や環境音声の再生が容易

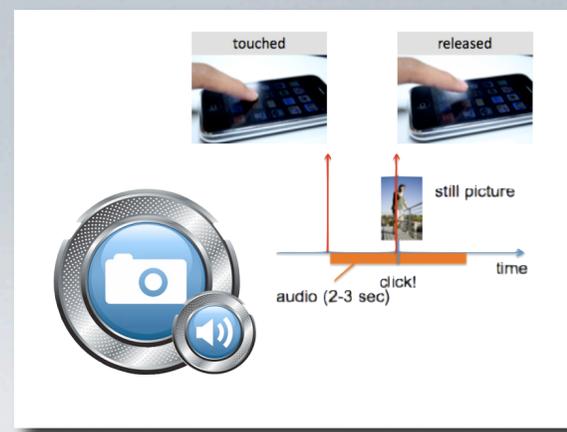
Exif Sub IFD

- Exposure Time (1 / Shutter Speed) = 3018/100000 second = 1/33.13453 second = 0.03018 second

AudiblePhoto EXIF Data

Exif GPS IFD

- GPS Version ID = 0x02,0x02,0x00,0x00
- GPS Latitude Reference = N
- GPS Latitude = 35/1,39/1,4940/100 [degrees, minutes, seconds] ==> 35° 39' 49.4"
- GPS Longitude Reference = E
- GPS Longitude = 139/1,35/1,3645/100 [degrees, minutes, seconds] ==> 139° 35' 36.45"
- Links to online mapping websites:
 - o Google Maps
 - o Yahoo! Maps
 - o MSN Maps & Directions
 - o Mapquest
 - o Open KML data with Google Earth
 - o Save KML data to file
 - o Save KML data to file and open with Google Earth
- GPS Map Datum = WGS-84
- GPS Processing Method = 0x41,0x53,0x43,0x49,0x00,0x00,0x00,0x47,0x50,0x53,0x2d,0x46,0x49,0x58

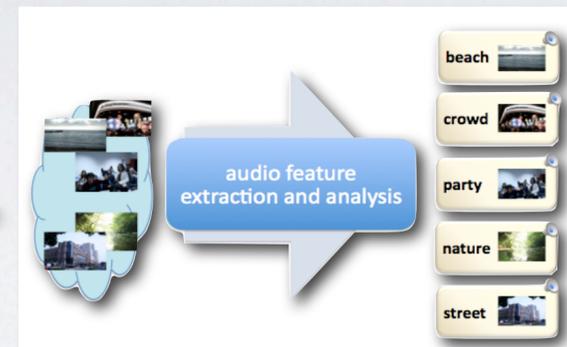


iPhoneアプリ、AudiblePhotoを開発

- 写真・音声キャプチャ・アプリケーションをiPhone 3G/3GSをプラットフォームとして開発
- iPhone上で音声・写真のハイブリッドデータを作成、ファイルをPC側に転送する際iPhone上でウェブサーバを構築

環境音声をベースにした写真の自動タギング

- 録音した環境音声の特徴点を抽出して、予め学習したレファレンスデータのクラスチャとマッチングする
- マッチング結果は写真のタグとして登録し、既存の技術に無い環境音声をベースにした写真のタグを実現



クライアント・サーバ型

- iPhone上で録音した音声をサーバに送る
- サーバ側で、音声特徴点抽出や自動タギングを行い、その結果をiPhoneに送り返す