

傳統模式 翻卡模式 雜誌模式 拼貼模式 側欄模式 快照模式 時間表模式

## 壓縮跟蹤Compressive Tracking

## 壓縮跟蹤Compressive Tracking

zouxy09@qq.com [mailto:zouxy09@qq.com]

<http://blog.csdn.net/zouxy09> [http://blog.csdn.net/zouxy09]

好了，學習瞭解了稀疏感知的理論知識後，終於可以來學習《Real-Time Compressive Tracking》這個paper介紹的感知跟蹤算法了。自己英文水平有限，理解難免出錯，還望各位不吝指正。

下面是這個算法的工程網站：裡面包含了上面這篇論文、Matlab和C++版本的代碼，還有測試數據、demo等。後面我再學習學習裡面的C++版本的代碼，具體見博客更新。

<http://www4.comp.polyu.edu.hk/~cslzhang/CT/CT.htm>

[<http://www4.comp.polyu.edu.hk/~cslzhang/CT/CT.htm>]

之前自己稍微學習了下稀疏感知的理論知識總結：

<http://blog.csdn.net/zouxy09/article/details/8118329>

[<http://blog.csdn.net/zouxy09/article/details/8118329>]

## 一、實時壓縮跟蹤：

感謝香港理工大學的Kaihua Zhang，這是他即將在ECCV 2012上出現的paper：Real-timeCompressive Tracking。這裡是他的介紹：

一種簡單高效地基於壓縮感知的跟蹤算法。首先利用符合壓縮感知RIP條件的隨機感知矩對多尺度圖像特徵進行降維，然後在降維後的特徵上採用簡單的樸素貝葉斯分類器進行分類。該跟蹤算法非常簡單，但是實驗結果很魯棒，速度大概能到達40幀/秒。

實際上，感覺上面這幾句話的介紹已經高度的概括了這個論文的主要思想。和一般的模式分類架構一樣：先提取圖像的特徵，再通過分類器對其分類，不同在於這裡特徵提取採用壓縮感知，分類器採用樸素貝葉斯。然後每幀通過在線學習更新分類器。當然，裡面還包含著很多細節的推導和優化了，下面我們從論文中一起來學習一下。

神經網絡編程入門

圖像處理和計算機視覺中...

視覺跟蹤綜述

計算機視覺領域的一些牛...

Learning中的代數結構的...

和機器學習和計算機視覺...

機器學習知識點學習

壓縮感知 (Compressive ...

壓縮感知 (Compressive ...

壓縮跟蹤Compres...

1

壓縮跟蹤Compressive Tr...

用單張2D圖像重構3D場景

目標檢測的圖像特徵提取...

目標檢測的圖像特徵提取...

目標檢測的圖像特徵提取...  
TLD (Tracking-Learning-...計算機視覺目標檢測的框...  
TLD視覺跟蹤算法淺析人臉檢測之Haar分類...  
簡單粗糙的指尖檢測方法...交互系統的構建之 (四) ...  
計算機視覺、機器學習相...交互系統的構建之 (三) ...  
光流Optical Flow介紹與O...交互系統的構建之 (一) ...  
圖像分割之 (四) OpenC...交互系統的構建之 (二) ...  
圖像分割之 (三) 從Grap...

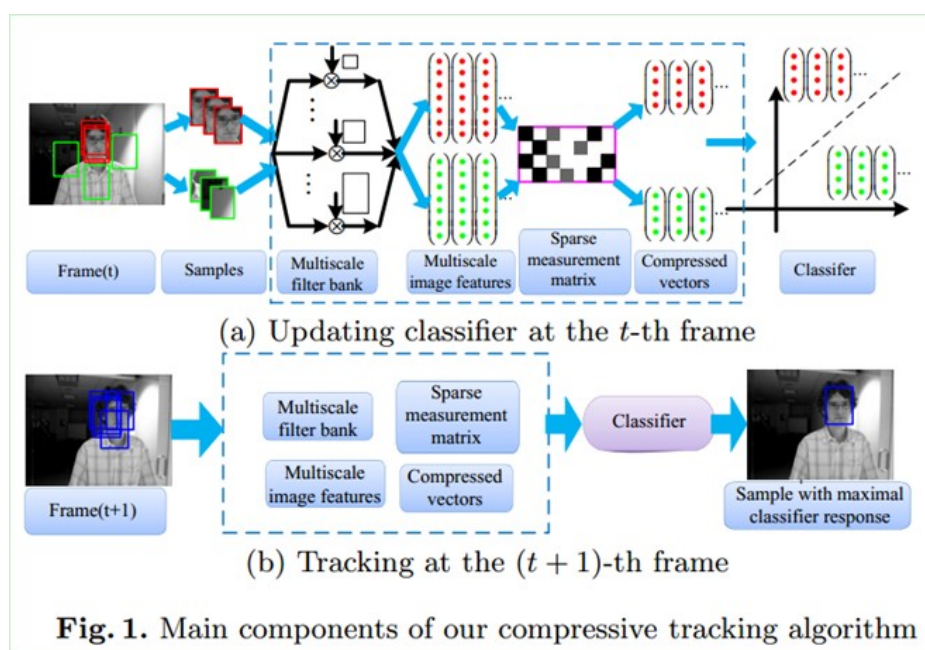
傳統模式 翻卡模式 雜誌模式 拼貼模式 側欄模式 快照模式 時間表模式

|   |
|---|
| 神經網絡編程入門                                  |
| 圖像處理和計算機視覺中...                            |
| 視覺跟蹤綜述                                    |
| 計算機視覺領域的一些牛...                            |
| Learning中的代數結構的...                        |
| 和機器學習和計算機視覺...                            |
| 機器學習知識點學習                                 |
| 壓縮感知 (Compressive ...                     |
| 壓縮感知 (Compressive ...                     |
| 壓縮跟蹤Compres... 1                          |
| 壓縮跟蹤Compressive Tr...                     |
| 用單張2D圖像重構3D場景                             |
| 目標檢測的圖像特徵提取...                            |
| 目標檢測的圖像特徵提取...                            |
| 目標檢測的圖像特徵提取... TLD (Tracking-Learning-... |
| 計算機視覺目標檢測的框... TLD視覺跟蹤算法                  |
| 淺析人臉檢測之Haar分類... 簡單粗糙的指尖檢測方法...           |
| 交互系統的構建之 (四) ... 計算機視覺、機器學習相...           |
| 交互系統的構建之 (三) ... 光流Optical Flow介紹與O...    |
| 交互系統的構建之 (一) ... 圖像分割之 (四) OpenC...       |
| 交互系統的構建之 (二) ... 圖像分割之 (三) 從Grap...       |

## 二、主要思想：

我再囉嗦一下：通過稀疏感知理論可以知道，我們通過一個滿足RIP條件的非常稀疏的測量矩陣對原圖像特徵空間做投影，就可以得到一個低維壓縮子空間。低維壓縮子空間可以很好的保留高維圖像特徵空間的信息。所以我們通過稀疏測量矩陣去提取前景目標和背景的特徵，作為在線學習更新分類器的正樣本和負樣本，然後使用該樸素貝葉斯分類器去分類下一幀圖像的目標待測圖像片（感知空間下）。

## 三、具體工作過程如下：



(1) 在 $t$ 幀的時候，我們採樣得到若干張目標（正樣本）和背景（負樣本）的圖像片，然後對他們進行多尺度變換，再通過一個稀疏測量矩陣對多尺度圖像特徵進行降維，然後通過降維後的特徵（包括目標和背景，屬二分類問題）去訓練樸素貝葉斯分類器。

(2) 在 $t+1$ 幀的時候，我們在上一幀跟蹤到的目標位置的周圍採樣 $n$ 個掃描窗口（避免去掃描整幅圖像），通過同樣的稀疏測量矩陣對其降維，提取特徵，然後用第 $t$ 幀訓練好的樸素貝葉斯分類器進行分類，分類分數最大的窗口就認為是目標窗口。這樣就實現了從 $t$ 幀到 $t+1$ 幀的目標跟蹤。

傳統模式 翻卡模式 雜誌模式 拼貼模式 側欄模式 快照模式 時間表模式

|   |
|---|
| 神經網絡編程入門                                  |
| 圖像處理和計算機視覺中...                            |
| 視覺跟蹤綜述                                    |
| 計算機視覺領域的一些牛...                            |
| Learning中的代數結構的...                        |
| 和機器學習和計算機視覺...                            |
| 機器學習知識點學習                                 |
| 壓縮感知 (Compressive ...                     |
| 壓縮感知 (Compressive ...                     |
| 壓縮跟蹤Compres... 1                          |
| 壓縮跟蹤Compressive Tr...                     |
| 用單張2D圖像重構3D場景                             |
| 目標檢測的圖像特徵提取...                            |
| 目標檢測的圖像特徵提取...                            |
| 目標檢測的圖像特徵提取... TLD (Tracking-Learning-... |
| 計算機視覺目標檢測的框... TLD視覺跟蹤算法                  |
| 淺析人臉檢測之Haar分類... 簡單粗糙的指尖檢測方法...           |
| 交互系統的構建之 (四) ... 計算機視覺、機器學習相...           |
| 交互系統的構建之 (三) ... 光流Optical Flow介紹與O...    |
| 交互系統的構建之 (一) ... 圖像分割之 (四) OpenC...       |
| 交互系統的構建之 (二) ... 圖像分割之 (三) 從Grap...       |

$\mathbf{x}$  ( $m$ 維) 變換到一個低維的空間 $\mathbf{v}$  ( $n$ 維)，數學表達就是： $\mathbf{v} = \mathbf{R} \mathbf{x}$

在這裡 $n$ 遠遠小於 $m$ （這樣才叫降維嘛）。最理想的情況，我們當然希望低維的 $\mathbf{v}$ 可以完全的保留高維的 $\mathbf{x}$ 的信息，或者說保持原始空間中各樣本 $\mathbf{x}$ 的距離關係，這樣在低維空間進行分類才有意義。

Johnson-Lindenstrauss推論表明：可以隨機選擇一個適當的高維子空間（當然，需要比原始空間維度小），原始空間兩點的距離投影到這個子空間，能高概率的保留這種距離關係。（ $K+1$ 次測量足以精確復原 $N$ 維空間的 $K$ -稀疏信號）。

而Baraniuk證明了滿足Johnson-Lindenstrauss推論的隨機矩陣同樣滿足壓縮感知理論中的 restricted isometry property (RIP) 條件。所以，如果隨機矩陣 $\mathbf{R}$ 滿足Johnson-Lindenstrauss推論，那麼如果 $\mathbf{x}$ 是可壓縮的（或者說是稀疏的），我們就可以通過最小化誤差來從 $\mathbf{v}$ 中高概率恢復 $\mathbf{x}$ 。

所以原文就找到了一個非常稀疏的投影矩陣，不但滿足Johnson-Lindenstrauss推論，而且可以高效的實時計算。

## 4.2、隨機測量矩陣：

一個比較典型的滿足RIP條件的測量矩陣是隨機高斯矩陣，矩陣元素滿足 $N(0,1)$ 分佈。但是，如果 $m$ 的維數比較大的話，這個矩陣還是比較稠密的，它的運算和存儲消耗還是比較大的。而在原文，採用了一個非常稀疏的隨機測量矩陣，其矩陣元素定義為：

$$r_{ij} = \sqrt{s} \times \begin{cases} 1 & \text{with probability } \frac{1}{2s} \\ 0 & \text{with probability } 1 - \frac{1}{s} \\ -1 & \text{with probability } \frac{1}{2s} \end{cases} \quad (2)$$

Achlioptas證明了，上式 $s$ 取2或者3時，矩陣就滿足Johnson-Lindenstrauss推論。這個矩陣非常容易計算，因為它只需要一個均勻隨機數發生器就行，而且當 $s=3$ 時，這個矩陣非常稀疏，計算量將會減少2/3。如果 $s=3$ ，那麼矩陣元素有1/6的概率為1.732（表示根號3，懶得插入公式了），有1/6的概率為-1.732，有2/3的概率為0；



傳統模式 翻卡模式 雜誌模式 拼貼模式 側欄模式 快照模式 時間表模式

神經網絡編程入門

圖像處理和計算機視覺中...

視覺跟蹤綜述

計算機視覺領域的一些牛...

Learning中的代數結構的...

和機器學習和計算機視覺...

機器學習知識點學習

壓縮感知 (Compressive ...

壓縮感知 (Compressive ...

壓縮跟蹤Compres...

1

壓縮跟蹤Compressive Tr...

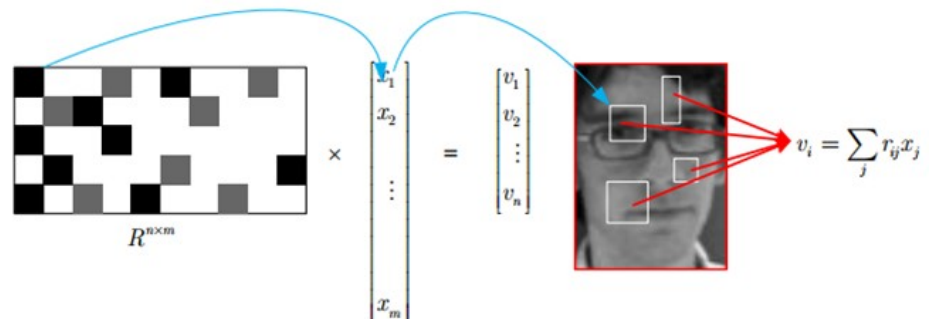
用單張2D圖像重構3D場景

目標檢測的圖像特徵提取...

目標檢測的圖像特徵提取...

目標檢測的圖像特徵提取...  
TLD (Tracking-Learning-...計算機視覺目標檢測的框...  
TLD視覺跟蹤算法淺析人臉檢測之Haar分類...  
簡單粗糙的指尖檢測方法...交互系統的構建之 (四) ...  
計算機視覺、機器學習相...交互系統的構建之 (三) ...  
光流Optical Flow介紹與O...交互系統的構建之 (一) ...  
圖像分割之 (四) OpenC...交互系統的構建之 (二) ...  
圖像分割之 (三) 從Grap...

## 4.3、提出的算法：



上圖表明一個 $n \times m$ 的稀疏矩陣，它可以將一個高維圖像空間的 $\mathbf{x}$  ( $m$ 維) 變換到一個低維的空間 $\mathbf{v}$  ( $n$ 維)，數學表達就是： $\mathbf{v} = \mathbf{R} \mathbf{x}$ ；

其中，矩陣 $\mathbf{R}$ 中，黑色、灰色和白色分別代表矩陣元素為負數、正數和零。藍色箭頭表示測量矩陣 $\mathbf{R}$ 的一行的一個非零元素感知 $\mathbf{x}$ 中的一個元素，等價於一個方形窗口濾波器和輸入圖像某一固定位置的灰度卷積。

為了實現尺度不變性，對每一個樣本 $\mathbf{z} \in \mathbb{R}^{w \times h}$ ，通過將其與一系列多尺度的矩形濾波器 $\{h_{1,1}, \dots, h_{w,h}\}$ 進行卷積，每一種尺度的矩形濾波器定義如下：

$$h_{i,j}(x,y) = \begin{cases} 1, & 1 \leq x \leq i, 1 \leq y \leq j \\ 0, & \text{otherwise} \end{cases}$$

式中， $i$ 和 $j$ 分別是矩形濾波器（模版）的寬和高。然後將濾波後的的圖像矩陣展成一個 $w \times h$ 維的列向量。再將這些列向量連接成一個非常高維（ $(w \times h)^2$ 維）的多尺度圖像特徵向量 $\mathbf{x} = (x_1, \dots, x_m)^T$ 。維數一般在10的6次方到10次方之間。

我們通過採用上面的稀疏隨機矩陣 $\mathbf{R}$ 將 $\mathbf{x}$ 投影到低維空間的 $\mathbf{v}$ 。這個隨機矩陣 $\mathbf{R}$ 只需要在程序啟動時計算一次，然後在跟蹤過程中保持不變。通過積分圖，我們可以高效的計算 $\mathbf{v}$ 。

## 4.4、低維壓縮特徵的分析：

傳統模式 翻卡模式 雜誌模式 拼貼模式 側欄模式 快照模式 時間表模式

|   |
|---|
| 神經網絡編程入門                                  |
| 圖像處理和計算機視覺中...                            |
| 視覺跟蹤綜述                                    |
| 計算機視覺領域的一些牛...                            |
| Learning中的代數結構的...                        |
| 和機器學習和計算機視覺...                            |
| 機器學習知識點學習                                 |
| 壓縮感知 (Compressive ...                     |
| 壓縮感知 (Compressive ...                     |
| 壓縮跟蹤Compres... 1                          |
| 壓縮跟蹤Compressive Tr...                     |
| 用單張2D圖像重構3D場景                             |
| 目標檢測的圖像特徵提取...                            |
| 目標檢測的圖像特徵提取...                            |
| 目標檢測的圖像特徵提取... TLD (Tracking-Learning-... |
| 計算機視覺目標檢測的框... TLD視覺跟蹤算法                  |
| 淺析人臉檢測之Haar分類... 簡單粗糙的指尖檢測方法...           |
| 交互系統的構建之(四)... 計算機視覺、機器學習相...             |
| 交互系統的構建之(三)... 光流Optical Flow介紹與O...      |
| 交互系統的構建之(一)... 圖像分割之(四) OpenC...          |
| 交互系統的構建之(二)... 圖像分割之(三) 從Grap...          |

的計算量比核心理論，時常力不從心。因此，本文通過稀疏測量矩陣對這些數目龐大的Haar-like特徵進行壓縮，稀疏感知理論保證了，壓縮後的特徵幾乎保留原有圖像的信息。因此，我們可以直接對壓縮空間裡面的投影特徵進行分類，而避免了維數災難。

#### 4.5、分類器構建和更新：

對每個樣本 $\mathbf{z}$ （ $m$ 維向量），它的低維表示是 $\mathbf{v}$ （ $n$ 維向量， $n$ 遠小於 $m$ ）。假定 $\mathbf{v}$ 中的各元素是獨立分佈的。可以通過朴素貝葉斯分類器來建模。

$$H(\mathbf{v}) = \log \left( \frac{\prod_{i=1}^n p(v_i|y=1)p(y=1)}{\prod_{i=1}^n p(v_i|y=0)p(y=0)} \right) = \sum_{i=1}^n \log \left( \frac{p(v_i|y=1)}{p(v_i|y=0)} \right), \quad (4)$$

其中， $y \in \{0,1\}$ 代表樣本標籤， $y=0$ 表示負樣本， $y=1$ 表示正樣本，假設兩個類的先驗概率相等。 $p(y=1)=p(y=0)=0.5$ 。Diaconis和Freedman證明了高維隨機向量的隨機投影幾乎都是高斯分佈的。因此，我們假定在分類器 $H(\mathbf{v})$ 中的條件概率 $p(v_i|y=1)$ 和 $p(v_i|y=0)$ 也屬於高斯分佈，並且可以用四個參數來描述：

$$(\mu_i^1, \sigma_i^1, \mu_i^0, \sigma_i^0) \quad p(v_i|y=1) \sim N(\mu_i^1, \sigma_i^1), \quad p(v_i|y=0) \sim N(\mu_i^0, \sigma_i^0).$$

上式中的四個參數會進行增量更新：

$$\begin{aligned} \mu_i^1 &\leftarrow \lambda \mu_i^1 + (1-\lambda) \mu^1 \\ \sigma_i^1 &\leftarrow \sqrt{\lambda(\sigma_i^1)^2 + (1-\lambda)(\sigma^1)^2 + \lambda(1-\lambda)(\mu_i^1 - \mu^1)^2}, \end{aligned} \quad (6)$$

式中，學習因子 $\lambda > 0$ ，

$$\sigma^1 = \sqrt{\frac{1}{n} \sum_{k=0|y=1}^{n-1} (v_i(k) - \mu^1)^2} \quad \mu^1 = \frac{1}{n} \sum_{k=0|y=1}^{n-1} v_i(k).$$

上式可以由**最大化似然估計**得到。

傳統模式 翻卡模式 雜誌模式 拼貼模式 側欄模式 快照模式 時間表模式

神經網絡編程入門

圖像處理和計算機視覺中...

視覺跟蹤綜述

計算機視覺領域的一些牛...

Learning中的代數結構的...

和機器學習和計算機視覺...

機器學習知識點學習

壓縮感知 (Compressive ...

壓縮感知 (Compressive ...

壓縮跟蹤Compres... 1

壓縮跟蹤Compressive Tr...

用單張2D圖像重構3D場景

目標檢測的圖像特徵提取...

目標檢測的圖像特徵提取...

目標檢測的圖像特徵提取...

計算機視覺目標檢測的框...

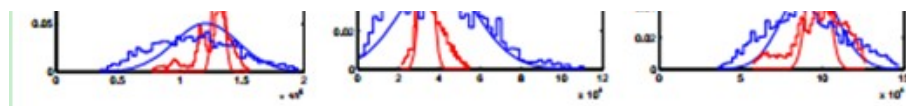
淺析人臉檢測之Haar分類...

交互系統的構建之 (四) ...

交互系統的構建之 (三) ...

交互系統的構建之 (一) ...

交互系統的構建之 (二) ...



圖中顯示了從某幀中的正樣本和負樣本提取出的三個不同特徵（低維空間下）的概率分佈。紅色和藍色階梯線分別代表正樣本和負樣本的直方圖。而紅色和藍色的曲線表示通過我們的增量更新模型得到的相應的分佈估計。圖說明了在投影空間，通過上式描述的在線更新的高斯分佈模型是特徵的一個良好估計。

## 五、壓縮跟蹤算法：

輸入：第 $t$ 幀圖像

1、在 $t-1$ 幀跟蹤到的目標位置 $l_{t-1}$ 的周圍（也就是滿足 $D^Y = \{z | ||l(z) - l_{t-1}|| < \gamma, \text{ 與 } l_{t-1} \text{ 距離小於 } \gamma\}$ ）採樣 $n$ 個圖像片，然後對這些圖像片進行特徵提取（降維），得到每個圖像片的特徵向量 $v$ 。

2、使用式（4）中分類器 $H(v)$ 對這些 $v$ 進行分類，找到最大分類分數的圖像片作為當前幀跟蹤到的目標，位置為 $l_t$ ；

3、採樣兩個樣本集： $D^\alpha = \{z | ||l(z) - l_t|| < \alpha\}$ 和  $D^\zeta, \beta = \{z | \zeta < ||l(z) - l_t|| < \beta\}$ 其中， $\alpha < \zeta < \beta$ ；

4、提取上述兩個樣本集的特徵，通過式（6）來更新分類器參數。

輸出：跟蹤到的目標位置 $l_t$ 和更新後的分類器參數。

張貼時間：2nd March 2014，張貼者：[alexchiang](#)

檢視留言