Detection of forged handwriting using wavelets

Prasad Chapekar (10407001),Patel Aaqib Ashfaq (10407603), Omkar Gune(10307062),
Badri Narayan Partho (09307903), Mustafa Saifee (07D07004)

1 Introduction

Wavelets were first used by Beverly Lytle and Caroline Yang to detect forged handwriting(reference). The concept is an adaptation of the procedure used by Simoncelli (reference) to detect art forgeries. A persons handwriting has certain distinguishing characteristics which include amount of curve of letters, slant, peculiar modifications to certain letters etc. Wavelets can capture the edge information, which often contains the characteristics of the writer. Wavelet transform of an image consists of 4 images, namely the first level approximation and incremental information when horizontal, vertical and diagonal filters are applied. The Wavelet transforms of different handwriting samples will consequently have different properties. Forgeries are expected to have different statistical parameters as compared to the actual handwriting. This is an effort to capture those differences by studying the statistical properties of wavelet coefficients and thereby distinguish between the two.

2 Steps involved in Detection

The detection procedure is broadly divided into 3 steps wherein the image is first divided into 8 parts and then each part is two level wavelet decomposed into 7 subbands. For every pixel of a high pass subband of the lower level (LH_1, HL_1 and HH_1 as shown in Figure), using some rule, 9 neighbours are extracted. Each pixel is approximated as a linear combination of these neighbouring pixels. Error is calulated between the actual value of the pixel and the approximated value. For each subimage, the co-effecients of the linear combination of each pixel of the high pass subbands and the corresponding error vector which is defined in following subsections the skewness is calculated for all these vectors. The same process is repeated for other image and then comparison is done using the Analysis of variance method or the ANOVA test. The following subsections detail out the aforementioned steps.

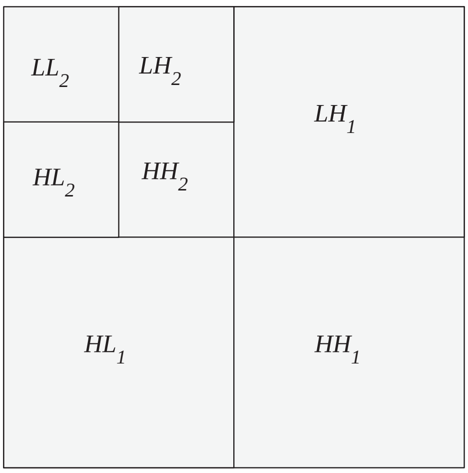

Figure 1: 2 level wavelet decomposition. Subbands are indicated

2.1 Image Division and Wavelet decomposition

Each image is broken into smaller part each of which is referred as a subimage. Since we are dealing with forged handwriting the images are divided along the columns. The division is done fixed number of parts for each image. Thereafter each subimage is 2 level wavelet decomposed [2] using the "Daubechies wavelet" of lenght 4, to obtain seven subbands LL_2, LH_2, HL_2, HH_2, LH_1, HL_1, HH_1 as shown in the Figure. The high pass subbands contain the necessary edge information and are hence very useful is detection processes. The overall edge information is a feature of an image which is in general varies with different images. Hence use of wavelet is justified in this detection process as different handwritings are expected to have different edge information. This is especially true for first level high pass subbands.

2.2 Neighbour Selection and Linear Prediction

Since we are interested in the first level high pass subbands we select neighbours of pixels in these subbands. The rule for selecting the neighbouring pixels depends upon the location of the pixel in consideration. If p denotes the pixel we are considering then we consider the 9 neighbours. In each of these cases 4 neighbours of p are n_{ip} for $i = 1, 2, 3, 4$ which are the immediate neighbours that share a common edge with the pixels. The rest of the neighbours are selected as shown below [1].

1. When the pixel $h \in LH_1$ and $h \equiv (x, y)$. In this case the following are the additional 5 neighbours
   - Vertical Cousin $v_h \in HL_1$ which has coordinates $(x + \frac{N}{2}, y - \frac{N}{2})$
   - Diagonal cousin $d_h \in HH_1$ which has coordinates $(x + \frac{N}{2}, y)$
   - Parent $P_h \in HL_2$ which has coordinates $(\frac{x}{2}, \frac{y}{2})$
   - Vertical Aunt $V_h \in HL_2$ which has coordinates $(\frac{x}{2} + \frac{N}{4}, \frac{y}{2} - \frac{N}{4})$
   - Diagonal Aunt $D_h \in HH_2$ which has coordinates $(\frac{x}{2} - \frac{N}{4}, \frac{y}{2}$
2. When the pixel $v \in HL_1$ and $v \equiv (x, y)$. In this case the following are the additional 5 neighbours
   - Diagonal Cousin $d_v \in HH_1$ which has coordinates $(x, y + \frac{N}{2})$
   - Horizontal cousin $h_v \in LH_1$ which has coordinates $(x - \frac{N}{2}, y + \frac{N}{2})$
   - Parent $P_v \in LH_2$ which has coordinates $(\frac{x}{2}, \frac{y}{2})$
   - Diagonal Aunt $D_v \in HH_2$ which has coordinates $(\frac{x}{2}, \frac{y}{2} + \frac{N}{4})$
   - Horizontal Aunt $H_v \in LH_2$ which has coordinates $(\frac{x}{2} - \frac{N}{4}, \frac{y}{2} + \frac{N}{4})$
3. When the pixel $d \in HH_1$ and $d \equiv (x, y)$. In this case the following are the additional 5 neighbours
   - Vertical Cousin $v_d \in HL_1$ which has coordinates $(x, y - \frac{N}{2})$
   - Horizontal cousin $h_d \in LH_1$ which has coordinates $(x - \frac{N}{2}, y)$
   - Parent $P_d \in HL_2$ which has coordinates $(\frac{x}{2}, \frac{y}{2})$
   - Vertical Aunt $V_d \in HL_2$ which has coordinates $(\frac{x}{2}, \frac{y}{2} - \frac{N}{4})$
   - Horizontal Aunt $H_d \in LH_2$ which has coordinates $(\frac{x}{2} - \frac{N}{4}, \frac{y}{2})$

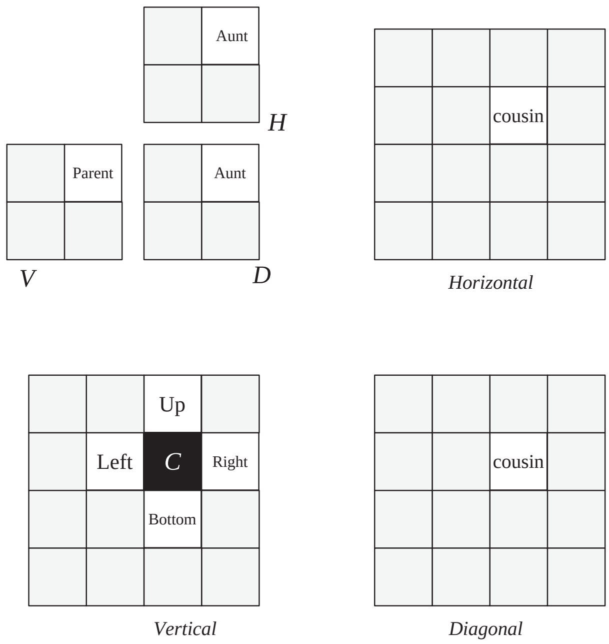

Figure 2: Neighbour selection for a pixel of the vertical HL_1 subband

Here N is the number of pixels in one direction. The images are taken such that after division they have become square subimages. The neighbours are chosen such that they relate most to the pixel of the given pixel. This is essential as we will be using these neighbouring pixel values to predict the given pixel value taking a linear combination of these 9 values. We have the following three predictions for $\{v, d, h\} \in \{LH_1, HL_1, HH_1\}$

$$\begin{aligned} v &= \sum_{i=1}^{4} w_{iv} n_{iv} + w_{5v} d_v + w_{6v} h_v + w_{7v} P_v + w_{8v} D_v + w_{9v} H_v \\ d &= \sum_{i=1}^{4} w_{id} n_{id} + w_{5d} v_d + w_{6d} h_d + w_{7d} P_d + w_{8d} V_d + w_{9d} H_d \\ h &= \sum_{i=1}^{4} w_{ih} n_{ih} + w_{5h} v_h + w_{6h} d_h + w_{7h} P_h + w_{8h} V_h + w_{9h} D_h \end{aligned}$$

The above equations can be conviniently written in the Matrix form as follows

$$\vec{V} = Q_v \vec{W}_v$$
$$\vec{D} = Q_d \vec{W}_d$$
$$\vec{H} = Q_h \vec{W}_h$$

The vectors $\vec{V}$, $\vec{D}$ and $\vec{H}$ are collection of the predicted pixel values for pixels belonging to LH_1, HL_1 and HH_1 respectively. Q_v, Q_d and Q_h denote the neighbouring pixel matrix where i^{th} row is set of neighbours for the i^{th} pixel or element of $\vec{V}$, $\vec{D}$ and $\vec{H}$ respectively. W_v, W_d and W_h denote the respective weights of the linear combination. It is easy to see that there may not be a unique set of values of any of W_v, W_d and W_h that solve the above system of equations. In fact there may not be any solution at all. Hence we solve these equations in the least square sense i.e. such that the final solution will have the least square error when compared for various pixels. Once we get the weight vectors we calculate the error vectors for each subband pixel. The error vector is defined for each of the three cases as

$$\begin{aligned} \vec{E_v} &= \log_2(|\vec{V}|) - \log_2(|Q_v \vec{W}_v|) \\ \vec{E_d} &= \log_2(|\vec{D}|) - \log_2(|Q_d \vec{W}_d|) \\ \vec{E_h} &= \log_2(|\vec{H}|) - \log_2(|Q_h \vec{W}_h|) \end{aligned}$$

We now have with us 6 vectors for each subimage, viz. $\vec{W_v}$, $\vec{W_d}$, $\vec{W_h}$, $\vec{E_v}$, $\vec{E_d}$ and $\vec{E_h}$ which charactarize each subimage of an image. If we have N subimages of an image then for each image there would be $6N$ such vectors. For each vector we calculate skewness which is the measure of the assymetry of the elements in a particular vector. A positive value of skewness indicates that the first few elements in the vector are larger than the latter ones and vice versa. The skewness of a vector $\vec{X}$ whose elements x_i have a mean value $\bar{x}$ and variance σ_x^2.

$$S(\vec{X}) = \sum_{i=1}^{N} \frac{(x_i - \bar{x})^3}{\sigma_x^3}$$

A matrix $I_{4\times N}$ representing the image in terms of these skewness values is constructed by arranging the rows such that $S(\vec{W_v})$ of each subimage occupy the first row, the second row is occupied by $S(\vec{W_d})$ for each subimage and then similarly 4 other rows are constructed. A similar matrix is constructed for the image to be constructed repeating the above procedure. Denote these matrices as I_1 and I_2 for the two images. Each row of the matrix I_1 and I_2 are passed through ANOVA test for comparison. Carefully choosing a threshold value for the test results in proper detection of forgery. The ANOVA test is discribed breifly in the following subsection

2.3 ANOVA Test

ANOVA stands for Analysis Of Variance. It is a test performed to test whether two or more groups of data come from the source of same mean. ANOVA test computes the F-statistic, a vital parameter which contains information about the similarity or difference between the different sources. The p-value, also obtained from this test, is derived from the cdf of F. This value is nothing but the probability that the groups come from the same source (or at least, source with the same mean). A large the value of p, indicates that it is more likely that the two or more group of readings were taken from the same source. Alternatively, a small value suggests that at least one group of data is significantly different than others. For detailed calculation of the F-statistics, refer [3]

3 Simulation and Results

The simulation was done by breaking the images to be compared into 8 parts along the column. The process described above was perforemed on each image.

Simulation results show that the technique employed here performs reseanably well given its simplicity in terms of computation and approach. In our simulation, we get 6 values from the ANOVA test when we compare the six sets of skewness. The threshold in our case is kept 0.8, i.e. *if one or more of the six values are more than 0.8*, we conclude that both the handwritings are the same. We tested this threshold on 132 comparisons between different handwritings and fonts. Approximately 75% of the time (described below), the simulation made a correct decision. This can be further improved by taking other combinations of neighbours, or taking a higher level of wavelet decomposition. There are four possible outcomes that arise. The cases when the two handwritings (in the images)

- match and the result is a *match*
- do not match and the result is a *mismatch*
- match and the result is a *mismatch*
- do not match and the result is a *match*

The first two outcomes are called as correct outcomes. The number of correct outcomes are denoted as N_c. The third outcome is called as a miss and number of such outcomes is denoted by N_m. Finally the fourth outcome is referred to as false positives and the number of such outcomes is denoted by N_f. We now define the following to measures

- Precision

$$P = \frac{N_c}{N_c + N_f}$$

- Recall

$$R = \frac{N_c}{N_c + N_m}$$

These two measures in general provide a good information about the robustness in terms of the selected threshold for the algorithm. In the event that the P value is very less when compared to the R value it implies that the threshold we have chosen in our algorith (in our case ANOVA) is less than the optimal. A similar P and R value is indicative of optimal threshold selecetion.

The simulation had the following experimental results. Let N_T denote total number of comparisons

Table 1: Correct, Misses and False positives

Measure	Value
N_T	132
N_c	98
N_f	25
N_m	9
P	0.74
R	0.92

From the table it seems that the treshold selection is not optimal. However, the threshold optimality discussed earlier holds only when the number of matches and mismatches expected should be the same. In our simulation the data simulated had higher number of expected mismatched (107 to be precise) and lower number of mathces (25 to be precise). Hence the optimal threshold will slightly shift in favour of the precision value.

4 Conclusion

We have seen the ability of wavelets to distinguish between different handwritings. Precisely we have seen the edge charactarization property of the high pass subbands obtained after wavelet decomposition. The linear predictor is shown to be a good feature of an image in such applications. Moreover, the linear predicted value was used as a feature for a particular subimage. Skewness as a third order moment was used to measure of the assymetry of the values of the elements in the vectors. Finally, the ANOVA test was used to differentiate between different images based on the skewness values obtained for the different vectors. One can conclude with the simplicity of approach and less compuational complexity this method is robust to a good extent and can be made more robust by carefull selection of neighbours that relate very highly to the pixel.

References

[1] R.W. Buccigrossi and E.P. Simoncelli. “Image compression via joint statistical characterization in the wavelet domain”. *Image Processing, IEEE Transactions on*, 8(12):1688 –1701, December 1999.

[2] C. Lytle, B.; Yang. “Detecting forged handwriting with wavelets and statistics”. *Rose-Hulman Undergrad. Math. J.*, 2006.

[3] Wikipedia. Analysis of variance — Wikipedia, the free encyclopedia. `http://en.wikipedia.org/wiki/Analysis_of_variance`, 2011. [Online; accessed 15-February-2011].