The method described by Michael Jones and Paul Viola in their paper, Robust Real-Time Face Detection, is designed to process images and perform face detection extremely rapidly. This method uses three unique processes that allow for the speed of the facial detection. They are the Integral Image, feature classification, and feature combination. The first process, the Integral image, detects and computes portions of the image to find features quickly. The second is a method to determine unique features among a set of larger potential features to find the most likely locations for features. The third being a way to combine the features discovered in the second process in a way to allow the removal of background information not needed in the facial detection method described. The experiment that this paper is written over did a test on many images of faces, to test the method and found that the method was comparable accuracy to other methods of the time, and was faster by far.

This method is different from others at the time in the fact that it works not with color, or pigment in pixels of a picture, but can function with gray-scale images, not requiring multiple images. Using color or multiple images also increases the speed of the processing. The fore mentioned contributions from the method are the primary causes for this improvement. The Integral Image does not work with image intensity, instead performing computations on each pixel. After these short calculations are finished, features can be created in constant K time, regardless of the features’ size, or location in the image. The second process takes the large library of features obtained from the first process and picks out the important features, using learning algorithms with AdaBoost to classify features in different levels of importance. The third step is combining the different important features found in the second process together, forming more complex features, and thus using predictive analysis to find likely regions of the face for the detector to use next. These three different processes summarize the steps for facial recognition this paper discusses.

When these steps are performed, they focus on using features instead of pixels. This is one of the key differences between this form of facial recognition and earlier versions of software detection. The features are used instead because working with features is far faster than working with pixels. The features have three different forms, all of which examine rectangles in areas and compare them. The two-rectangle feature compares the sums of two different rectangles that are the same size adjacent to each other. The three-rectangle compare two rectangles that are adjacent to a center rectangle, and the sum of those two rectangles is subtracted from the center rectangle. With the fourth-rectangle, four rectangles are made, and the difference of the two sets of rectangles is computed. The way that these rectangle features are calculated quickly is by using the integral image, which is a recursive method to sum up all the different pixel data in a selected area. The integral image can be summed up such that the sum of the data can be expressed as an integration of the rows and columns of the image times the box that is being analyzed.

The next step taken, after working with features, which had been worked with before, is that not all features, which includes around 150 k rectangle features, can be reduced to include only a small amount of those features are necessary to create a classification for a face and be able to recognize it. Obviously, the challenge with this is to find the significant classifiers. This is where AdaBost is used to help with improving given learning algorithms, which is essentially what is needed to learn what these classifiers are. It does this by rejecting options (or in this case, features) with low classification, and combines the remaining classifications to find stronger classifications. However, as this does not have a high accuracy rate, sequences of learning problems are used to help analyze which of the remaining classifications are higher, and assign weight to them. At that point, classifications have weights to them, and the classification process starts over, with a “perception” of how classifications will turn out, leading to the process finding good accuracy rates with the discovered higher accuracies that are left.

When working to create the classifiers, an algorithm called *the Attentional Cascade* is an algorithm used to cascade classifier to improve detection accuracy and improve speed both by using smaller classifiers, which are more efficient, reject a greater amount of false classifications, and detect most positive classifiers. It does this by starting with two classifiers with an adjustable threshold for confirming classifier connections (to form stronger classifiers). This is always able to detect when a face is there, and only sees a face that isn’t there half the time there isn’t a face. This two-feature classifier requires a small amount of work, and with the high rejection rates, this is an accurate method to generate results.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Detector | 10 | 31 | 50 | 65 | 78 | 95 | 167 | 422 |
| Viola-Jones | 76.1% | 88.4% | 91.4% | 92.0% | 92.1% | 92.9% | 93.9% | 94.1% |
| Rowley-Baluja-Kanade | 83.2% | 86.0% | - | - | - | 89.2% | 90.1% | 89.9% |
| Schneiderman-Kanade | - | - | - | 94.4% | - | - | - | - |
| Roth-Yang-Ahuja | - | - | - | - | 94.8% | - | - | - |

This new method of facial detection that was discovered has far more accuracy than several other methods used at the time. The experiments done using over 4000 photographed faces confirm this, considering that the faces were not rotated to make the face have vertical symmetry (eyes at same level), as well as the fact that the overall process was 15 times faster than recently created methods as approximately as accurate, only changing within 5% of false positive face detection. Also, this study focuses on different options for different detection sensitivity, whereas other studies did not perform as comprehensive a list. In the table listed above, you can see the accuracy differences for different sensitivities, represented for different studies. Since these studies did not include this information, they likely did not take into account of the need to consider if the facial detection would need to be adjusted for particular circumstances. Because of this, this method definitely shows various new applications for facial detection, can likely be used in several different applications, and has been improved upon even further since.