Prime Vision

Author: Martijn R. Nuijt (Prime Vision)
Version: 1.3, March 2003

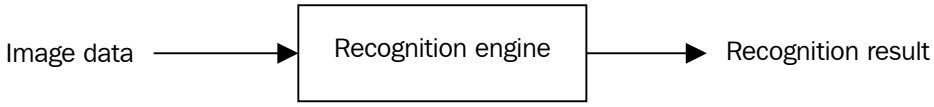
Basic recognition concepts

Introduction

This White Paper explains the basic concepts used in Prime Vision's recognition technology. The paper starts with a simple representation of a recognition module, before introducing the concept of thresholding. Finally, we explain how these concepts can be applied to optimise the costs of an automated process. The glossary at the end of this paper can be used as a reference.

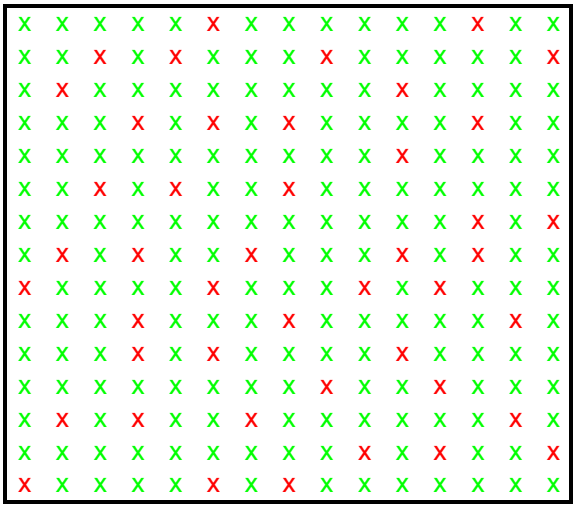
Basic recognition module

A basic recognition engine can be presented as follows.



The input for this engine is the image data of the item that needs to be recognised. The output is a recognition result: an interpretation into classes of the image data. The recognition result can be either correct or incorrect. The performance of the recognition engine is expressed by the *recognition rate*. This the percentage of items that are recognised correctly.

Suppose that we have an engine with a recognition rate of 70%. We can then represent the recognised items as follows:



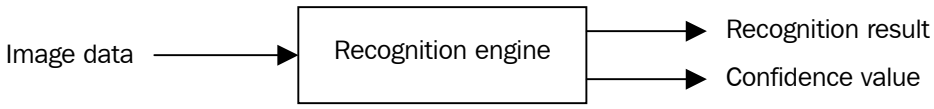
The green marks represent the items that were recognised correctly, the red marks the ones that were recognised incorrectly. Unfortunately, it is only possible to distinguish between these after verifying the recognition results.

Since, in this case, 30% of the results are incorrect, the recognition engine will not be very useful in practice if accurate results are required. In order to make better use of this performance, we would need to know in advance which items are recognised correctly and which are not.

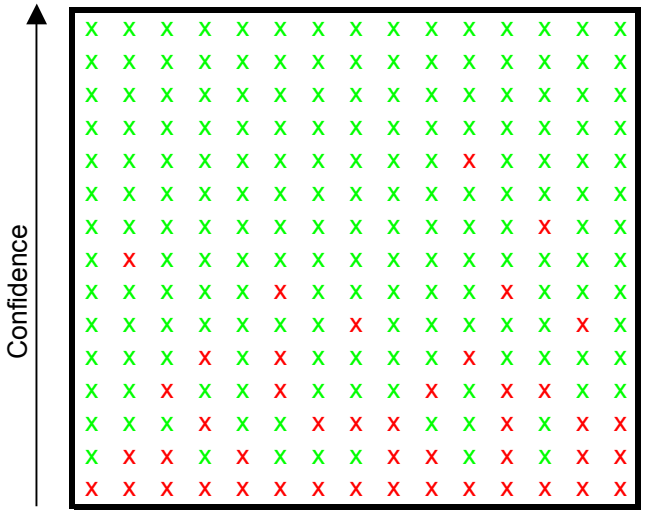
Introducing the confidence value

Only a perfect recogniser can indicate whether its result is correct or incorrect. However, good recognition algorithms will provide some kind of value indicating how reliable the result of the recognition is. This value is called the *confidence value*. A good recognition engine will provide a meaningful confidence level together with the recognition result.

A recognition engine can then be presented as follows.



Using the confidence value, we can sort the recognition results according to their reliability. The items from the previous figure can now be represented like this:

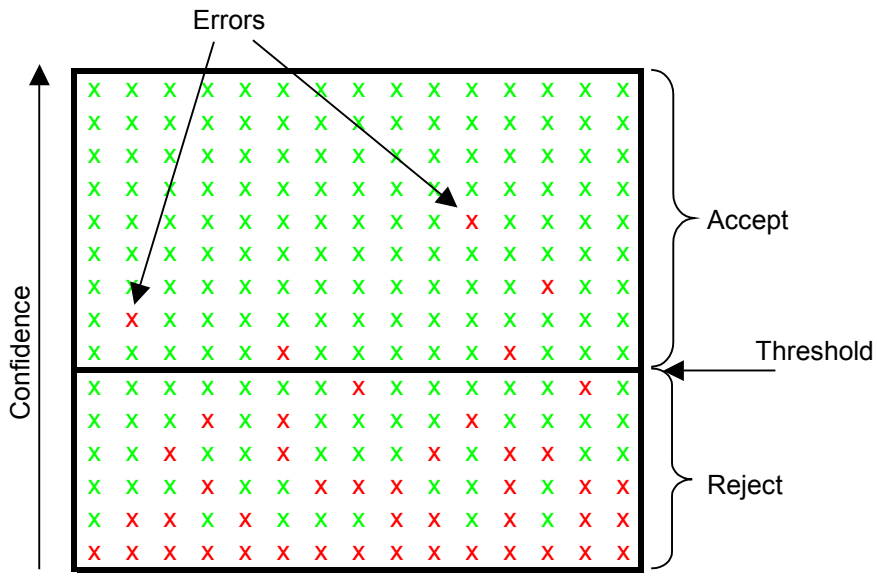


In the upper part side of this diagram, where the confidence is high, most of the items are recognised correctly; going down, the fraction of items recognised incorrectly slowly increases.

A higher confidence value indicates that the recognition result is more reliable; however, it is still no guarantee that the result is actually correct.

Thresholding

We can use the confidence value to get more reliable results. This is done by setting a threshold for the confidence value. Every recognition result with a confidence level above this threshold will be accepted. If the confidence level is below the threshold, it will be rejected. This is illustrated in the next figure.



The fraction of items with a confidence higher than the threshold is known as the *read rate* or *accept rate*. The fraction below the threshold is the *reject rate*. The accept rate plus the reject rate always add up to 100%. Within the read rate, there will probably still be some errors. The fraction of errors within the accepted items is called the *error rate*. Complementary to the error rate is the *accuracy*: the fraction of correctly recognised items within the read rate.

If the threshold is set to a high value, the read rate will be low, but the accuracy will be high. A lower threshold value will result in a higher read rate with a lower accuracy.

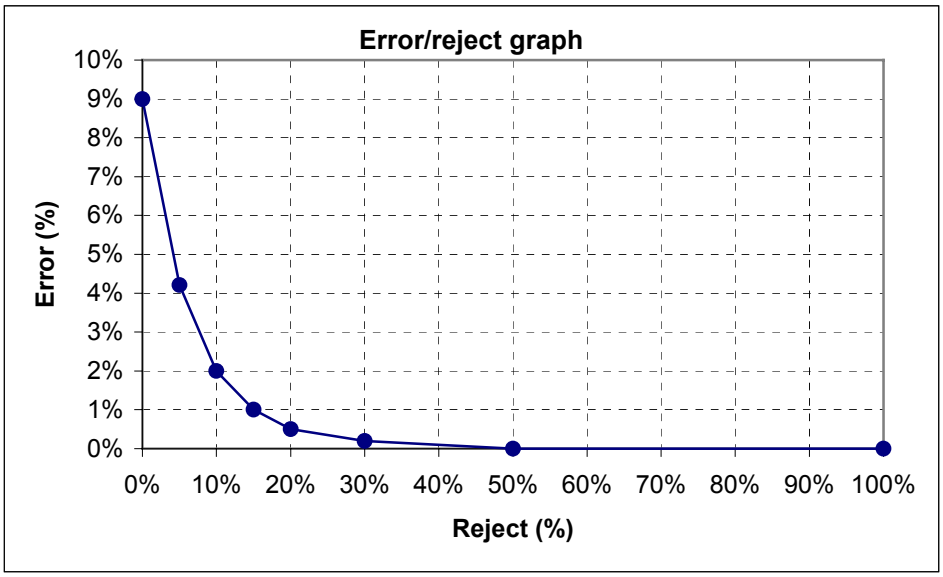
Note that it is meaningless to specify a read rate without specifying the accuracy or error rate. Also note that read rate and recognition rate are quite different concepts. The recognition rate is a fixed performance indicator, whereas the read rate can be adjusted to every possible value (obviously affecting the accuracy).

In many cases, the rejected items need to be handled by humans to obtain a reliable classification result. The accepted items, however, are recognised reliably enough to pass without any human inspection.

Understanding the error/reject graph

As explained above, the threshold value determines which combination of accuracy and read rate is achieved. Therefore, it is important to choose the right threshold value. The *error/reject graph* is used to represent all the necessary information for choosing the right confidence value.

An example of an error/reject graph is shown below.



In this graph, the error rate is plotted versus the reject rate. The bottom left corner of this graph represents the perfect recogniser: zero rejects with no errors. In other words, all items will be read correctly without any incorrect recognition or rejection. Going to the right, the reject rate increases; moving upwards in the graph, the error rate increases.

Looking at the example graph, we find that we can choose a reject rate of 20% with an error rate of 0.5%. In this case, the read rate is (100-20)%=80%. It is also possible to choose a read rate of 85% with a higher error rate of 1%. If we really want high accuracy, a read rate of 70% is possible with a 0.1% error rate (i.e. 99.9% accuracy). The graph shows that even a zero error rate is feasible if we are satisfied with a 50% read rate.

The graph also shows the recognition rate of this recogniser: at the point of zero rejects, an error rate of 9.0% is achieved. From this, we can conclude that the recognition rate of this recogniser is 91.0%.

Optimising efficiency

As we can see, the error/reject graph enables us to select different settings with different read/error rate characteristics. An important question at this point is: which of these settings is the best one, given the application? Generally speaking, the best setting is the one that optimises the total costs of the operational process. Below, we will explain a very straightforward way to look at this issue.

If we consider the costs of an operational process containing automated classification, we can distinguish between two cost factors:

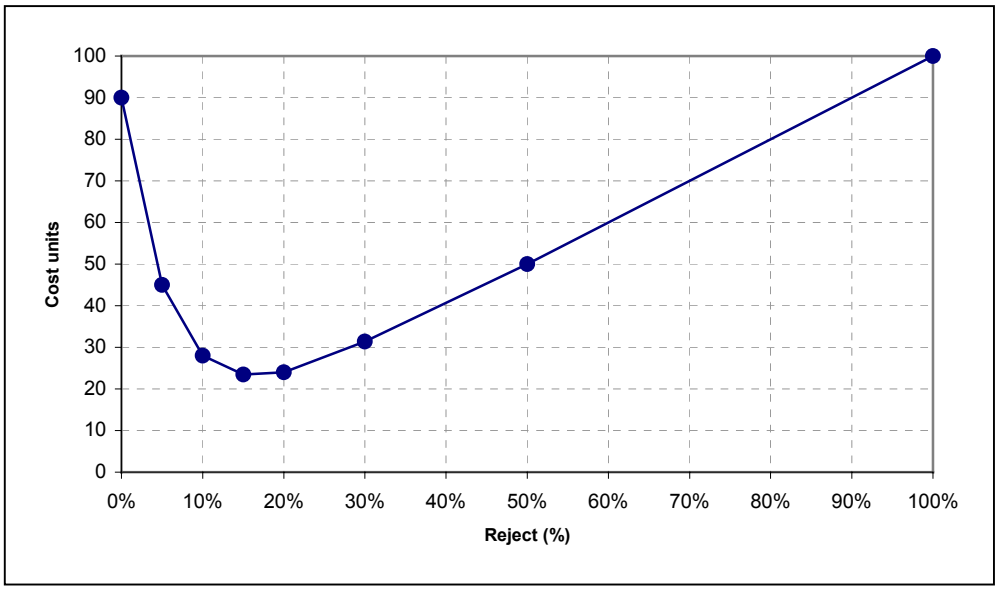
Costs of a reject — The extra costs as a result of an item being rejected; e.g. the costs of manually classifying the item.

Costs of an error — The extra costs as a result of an item being recognised incorrectly; e.g. the costs of correcting the error at a later stage in the process.

Obviously, the costs of an error will be higher than the costs of a reject. Costs due to errors tend to be more difficult to determine than costs due to rejects.

If the costs of rejects and errors are known, it is possible to calculate the costs of the process for each of the read/error rate settings.

As an example, the process costs are shown in a graph below for the error/reject curve presented earlier. In this case, we assume that the costs of an error are 10 times larger than the costs of a reject.



As the graph shows, the lowest costs are achieved at a read rate of approximately 85%. This corresponds to an error rate of 1.0%.

Glossary

Recognition rate	Fraction of items for which the recognition result is correct.
Confidence value	A value indicating the reliability of a recognition result. The higher the confidence value, the higher the probability that the recognition result is correct.
Threshold	Lowest confidence value for which a result is still accepted. The threshold needs to be chosen in such a way as to optimise the overall system performance.
Accept rate	Fraction of items for which the confidence value is above the threshold
Read rate	Synonym for Accept rate
Reject rate	Fraction of items for which the confidence value is below the threshold. Accept rate and reject rate add up to 100%.
Accuracy	Fraction of correctly recognised items within the accepted items
Error rate	Fraction of incorrectly recognised items within the accepted items. Accuracy and Error rate add up to 100%.