Document Analysis with Neural Net Circuits

Hans Peter Graf

AT&T Bell Laboratories, Holmdel, NJ 07733

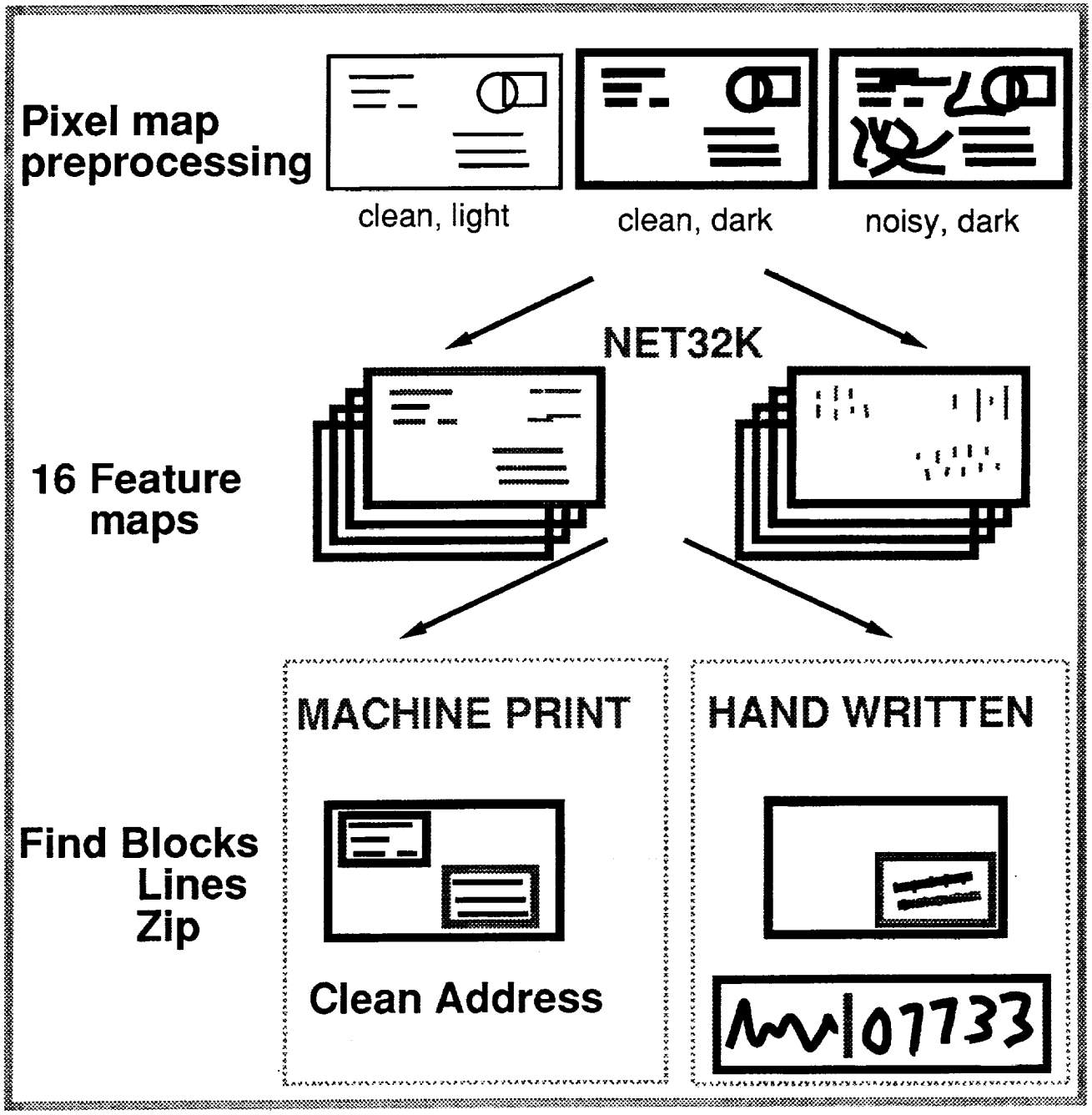
Document analysis is one of the main applications of machine vision today and offers great opportunities for neural net circuits. Despite more and more data processing with computers, the number of paper documents is still increasing rapidly. A fast translation of data from paper into electronic format is needed almost everywhere, and when done manually, this is a time consuming process. Markets range from small scanners for personal use to high-volume document analysis systems, such as address readers for the postal service or check processing systems for banks.

A major concern with present systems is the accuracy of the automatic interpretation. Systems tend to work well, if the image is not too complex and its quality is good, i.e. there is no noise in the image and the print quality is good. Todays algorithms, however, fail miserably when noise is present, when the print quality is poor or when the layout is complex. A common approach to circumvent these problems is, to restrict the variations of the documents handled by a system.

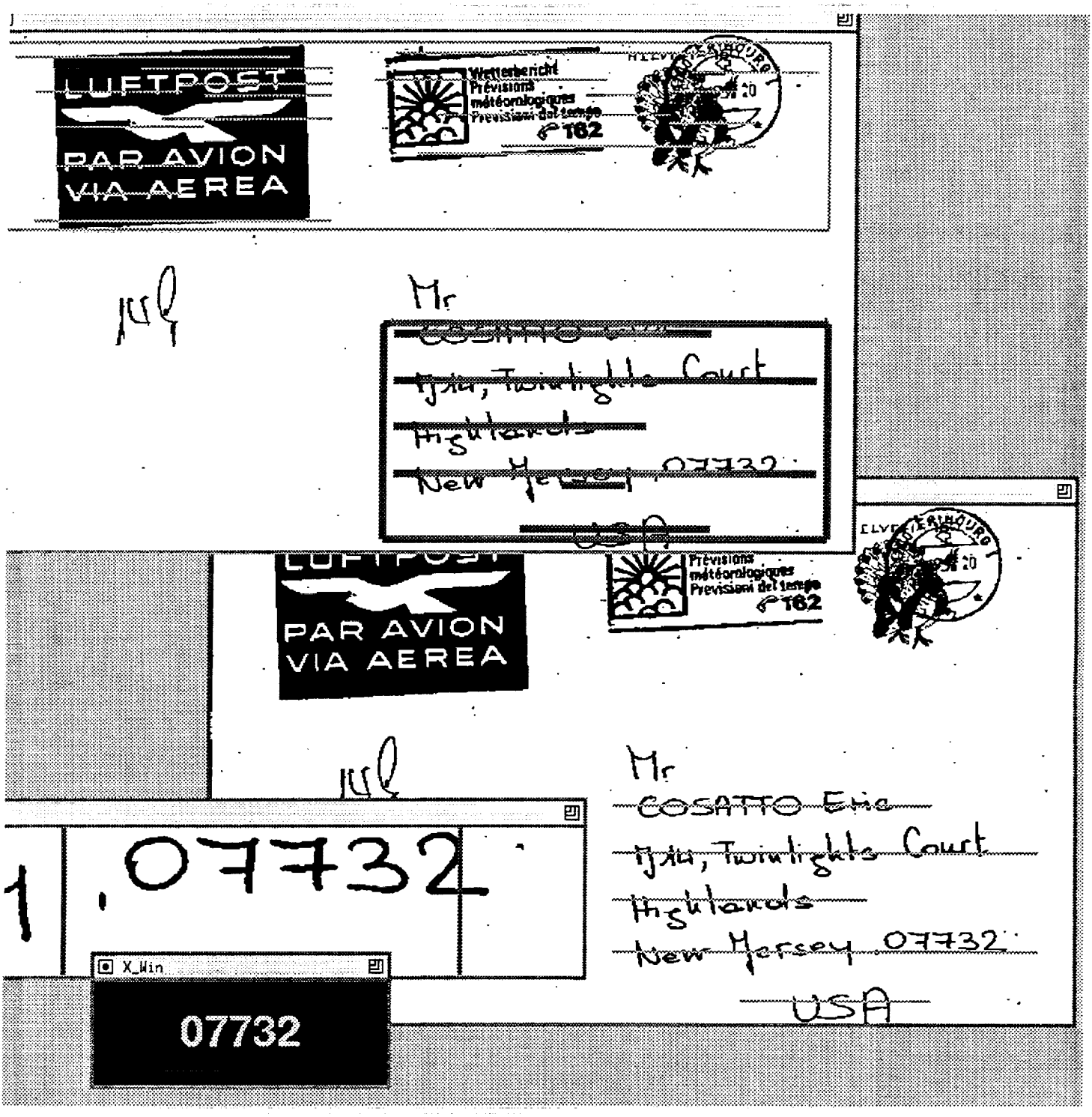
Improving the robustness of algorithms, to deal with a wider variety of conditions, seems always to lead to algorithms requiring an enormous amount of computation. This is a good opportunity for specialized circuits, such as neural net chips. Key for a successful integration of such a circuit into an application is that all the algorithms, from start to end, are taken into account and that the throughput of each stage is well balanced. Often neural net circuits were designed with one particular algorithm in mind, for example the character recognition. But in an application other processing steps, such as the layout analysis or just the discrimination between figures and text, may require more computation and represent the throughput bottleneck. It is clear by now that "pure neural network" solutions are suited for some aspects of document analysis, most notably the recognition of individual characters, but many problems are solved more effectively with other types of algorithms. The main problem for any hardware implementation is that algorithms applied in document analysis are still evolving and are changing rapidly. It is therefore easily possible that by the time a circuit is built and integrated into a system, newer algorithms with better performance and different compute requirements have been developed.

In our laboratory, we had the best luck with circuits implementing basic functions, such as convolutions, that can be used in many different algorithms. To illustrate the flexibility of this approach, three applications of the NET32K circuit are described: Locating address blocks, cleaning document images by removing noise, and locating areas of interest in personal checks to improve image compression. Several of the ideas realized in this circuit that were inspired by neural nets, such as analog computation with a low resolution, resulted in a chip that is well suited for real-world document analysis applications and that compares favorably with alternative, "conventional" circuits.

LOCATING ADDRESS BLOCKS

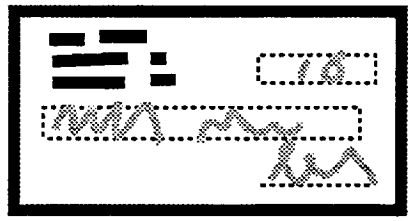


Announcing
FIX IT YOURSELF
The first how-to book
to make you a special movie

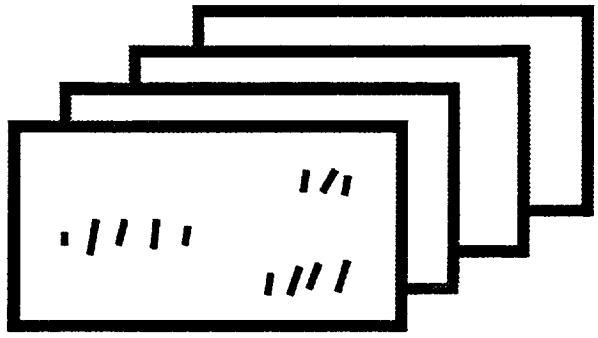

LUFTPOST
PAR AVION
VIA AEREA
Wetterbericht
Prévisions météorologiques
Previsioni del tempo
T62
Mr
Highlands
USA
LUFTPOST
PAR AVION
VIA AEREA
Prévisions météorologiques
Previsioni del tempo
T62
Mr
COSATTO Eric
Highlands
New Jersey 07732
USA
07732
X_Win
07732

Compress check image

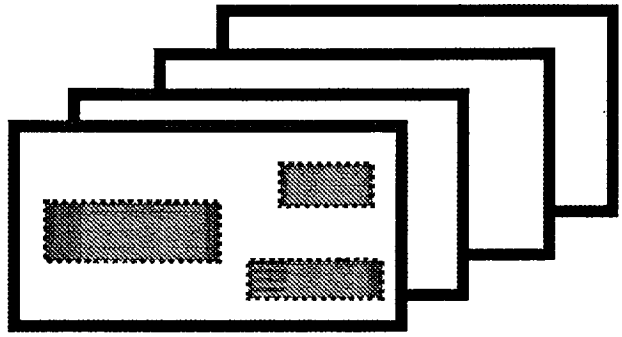
Check with information added by customer



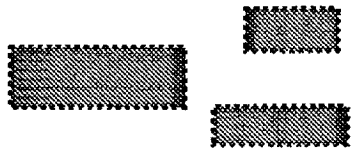
Extract features with NET32K: edges and strokes

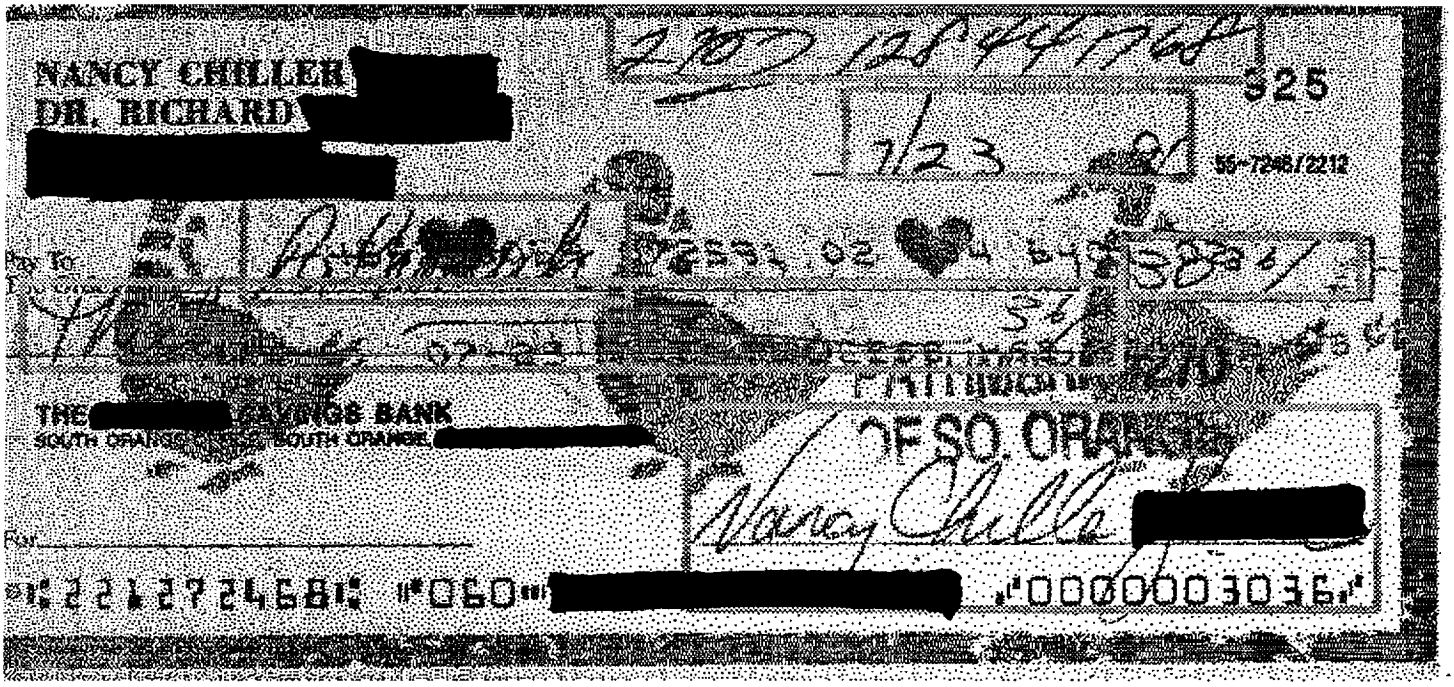


Identify areas with handwritten text.



Transmit only areas of interest.





(Portions of check obliterated to ensure privacy.)