DSP and FPGA based Implementation of a Video based Industrial Safety Application

JÖRG VELTEN, MARCO KRIPS, ANTON KUMMERT
Department of Electrical and Information Engineering
University of Wuppertal
42097 Wuppertal
GERMANY

Abstract: - Industrial safety devices are an interesting field of application for video surveillance and image processing. Complex production processes often can not be protected sufficiently by usual light row based safety devices. In the present paper, a fully video based security device is described. The device was developed for surveillance of press brakes, but it is easily adaptable to different types of machines or work places if several conditions are met.

Key-Words: - image processing, video based safety device

1 Introduction
The present paper describes a video based industrial safety application for protecting workers at press brakes. It can, nevertheless, also be adapted to other types of machines.

Press brakes are used to form prefabricated steel sheets. They are capable of bending sheet steel of 10 mm thickness whereas its pressing power amounts about 1500 t. It becomes obvious, that the worker who operates the press has to act strictly outside of dangerous areas to avoid serious injuries.

An often employed safety device for preventing accidents is the light curtain. It consists of several light barriers which are arranged in a column. The distance between a pair of light rows is small enough to detect any intruding object greater than a certain diameter. All light rows together define a planar area which can not be penetrated by these objects without being noticed.

Problems arise from complex-shaped workpieces which have to be guided by a worker during the pressing operation. Also the processing of small and long workpieces causes difficulties. Actually implemented solutions to these problems are compromises affecting security and/or ease of handling of the work piece. The latter justifies the demand for additional intelligence and scene understanding for the safety device.

2 Video based Industrial Safety Device

A plausible approach for realizing this intelligence and scene understanding is a video based surveillance of the work place. A video camera which is mounted above the press brake delivers top view images, like the one shown in figure 1, to a processing device. Scene understanding is obtained by applying image processing and recognition algorithms to these images. The processing device should be able to distinguish between workpiece and worker or safe and dangerous operation, respectively.
Extreme financial and speed constraints limit the choice of possible image processing algorithms. The price should be approximately the same as for usual light curtains which are simple structured and hence relatively cheap to construct. Furthermore, the whole processing should be done in approximately 20 ms.

The extremely small budget in combination with the restrictive time constraints requires the introduction of a-priori information to obtain at least partial scene understanding. The relevant information has to be minimized in order to meet the restriction with respect to implementation.

Succeeding steps of the course of implementation are:

- analysis of dangerous situation
- definition of required knowledge
- selection of suitable algorithms

3 Analysis of Dangerous Situations

For a better understanding of dangerous situations at press brakes, a short introduction to this kind of machine and its usual handling is shortly described. Press brakes mainly consist of three parts.

Upper tool: Usually the forceful part. It can be moved up- and downwards with a power of approximately 240 t. Movements can be performed in high or low speed. While high speed movements are required for positioning purposes, the full power of 240 t can only take effect in low speed.

Lower tool: Usually in fixed position. The upper tool presses the workpiece into the lower tool. Different shapes of bending can be realized by using different shaped lower and upper tools.

Limit stop: Used for fixing the workpiece. It limits the distance of the bending from one side.

For an effective workflow, different bendings of one workpiece are realized at different stations of one press brake. The whole length of the press brake (usually 2 up to 4 m) is subdivided into several positions. In this way, all necessary bendings of one workpiece can be executed at the same press brake by using several stations and by moving the limit stop.

The workflow can be described as follows:

1. The worker puts a raw, prefabricated workpiece into the press.
2. He uses a foot control to move the upper pressing tool with high speed into a pressing position. In this state, there is still a gap between upper tool and the workpiece, hence, the workpiece can still be positioned.
3. The worker presses the workpiece for correct positioning against the limit stop and initializes the bending operation by foot control.
4. The upper tool moves slowly, but with full power downwards and presses the workpiece into the lower tool until the workpiece is bent.
5. The upper tool moves upwards and releases the workpiece. If necessary, the limit stop changes its position for the next bending operation. During this step, the worker has to move the workpiece into the next bending position.

The further processing is continued at point 3 until bending of the workpiece is completed.

A closer look at this grading of the workflow reveals the most dangerous situations. Serious endangerment results from the second processing step, while the press brake is closed in high speed from a wide opened position. Even large objects can reach into the press and get almost cut by the downwards moving upper tool. Further injuries can happen during step 4 of the above described workflow. The gap between upper tool and workpiece or lower tool can still be large enough for fingers to fit in. If the pressing operation is accidentally initiated while the worker is fine positioning the workpiece, hands or fingers can be cut off by the slow but forceful movements of the upper tool. The second source of injury during the same processing step is crushing of hands between parts of the machine and the moving parts of the workpiece during its deformation. If for example the last side panel of a box is bent, the hand of the worker can be crushed between the already bent side panels and the machine. Also the moving limit stop can cause injuries during step 5 of the above mentioned workflow. If the limit stop is moving towards the lower tool, both together can act like shears, since there is almost no vertical gap between limit stop and tool. If the worker tries to move a canted workpiece by lifting it with his fingers behind the lower tool, he risks to be injured between limit stop and tool.

All these injuries can be prevented, if the hands are outside a “region of danger” during any of the above
mentioned situations, i.e. dangerous steps of the bending process. In order to guarantee this, the information about the exact positions of hands is required since they are the most endangered extremities.

4 Implementation concept

The strict budget constraints in connection with relatively high computational requirements lead to the necessity of considering all parts of the design simultaneously. Since this is a very complex task, it is advisable to subdivide the whole application into independent recognition tasks following a superior implementation concept.

- It is convenient to reduce the data word width as soon as possible. This reduces not only the computational complexity, but also the necessary hardware effort.

- The use of local instead of global operations is favorable, in particular in the first processing steps, when the data word length is relatively long.

Furthermore, iterative algorithms should be avoided, if the number of iterations is not predictable. For safety applications, the worst case computation time has to be taken into account for the hardware design which leads to a weak performance.

For the particular application of a video based safety device at press brakes the following basic ideas have been taken into account. The detection of hands is mainly based on analyzing the color information which allows to distinguish between hands and other objects. In other words, every pixel is classified whether it belongs to hands or not and the result can be coded by one bit. By doing this, the original color image is reduced to a binary image which represents a reduction of the data word length from 24 to 1 bit per pixel. Succeeding algorithms applied to binary images consider regional coherence of the scene. The output binary image of the classification block is finally compared to predefined masks which represent the different alarm regions. Any occurrence of pixels inside such a mask leads to an emergency alarm or a direct stop of the machine. Obviously, different arbitrarily shaped masks can be defined in order to cover

- different states of the workflow,
- different workpieces, and
- different stations of the press.

These ideas meet excellently the above described implementation concept. Since any pixel is classified individually, the storage of additional information from preceding images or the neighborhood can be avoided. Obviously, the mentioned ideas require, that detection quality is relatively high and robust against any kind of perturbation. Otherwise, a failure of the system could cause accidents.

5 Algorithms

As mentioned above, hand detection is based on color information. Already a simple artificial neural network provides good results for this task. A single perceptron (cf. [1]) is capable of dividing the feature space, spanned by usual workplace images, into classes “hand” and “not hand”. Figure 2 shows the feature space spanned by such an image. The perceptron calculates

\[ \text{Figure 2: The feature space spanned by a usual workplace image. Black dots represent hand pixel colors, gray dots represent other colors.} \]

the weighted sum of its inputs and applies a threshold to its result. Thus, the perceptron represents a simple threshold method. The advantage can be seen in the fact, that its well known learning rules can be realized by using DSPs or microprocessors. Consequently, the application can be re-trained if the detection results become too weak. Further investigations show (cf. [2]), that improved detection results can be obtained by using an artificial neural net consisting of 4 neurons and applying a modified backpropagation algorithm.

The output of the neural net stage still comprises misclassified pixels which are removed by exploiting the regional coherence of the image. Singular pixels appearing far apart from homogen regions are treated as
misclassified and can be removed by using mathematical morphological operations (cf. [3, 4]).

The resulting images show white pixels at hand locations and black ones everywhere else. In succeeding steps these binary images are compared to predefined masks representing alarm and danger regions. Any appearance of pixels inside such an area initiates a warning message or a complete stop of the machine.

Additional logic verifies via time and size, if the hands are detectable or not. Additional feature fusion steps based on the analysis of temporal and spacial continuity properties improve detection probability. If no hands can be localized, even after re-training of the algorithms, the worker is forced to wear colored gloves which leads to a significant robustness of the behaviour. Further information can be found in [5].

6 Implementation

The proposed algorithm has been implemented on two different hardware-platforms. For a first approach, it was programmed in C and implemented on a multipurpose board based on the TMS320C80 operating at 60 MHz tested under usual industrial environment conditions. The board was programmed using an image processing library. The processing time for one frame amounts to approximately 50 ms including the image acquisition. This value can be reduced significantly by using FPGAs for implementation of the perceptron and the morphological operations. Since pixels can be processed in parallel, pipelining leads to much faster results. While a final FPGA hardware-realization is still outstanding, simulations of VHDL-based implementations of the outlined processing structure lead to promising results. In connection with high speed cameras, sampling-frequencies of up to 100 MHz are processable by using low cost FPGAs. The whole design of a video-based safety device, including additional logic for supervision of the safety-device fits into one XILINX Spartan2 XC2S200 FPGA.

7 Conclusion

Careful analysis of the workflow at press brakes delivers a-priori information which permits the implementation of a video based safety device. In particular, this strategy allows a low-cost realization of the safety device. The algorithm presented has been proven to be very efficient with respect to hardware effort. Thus, it is an ideal component for implementation of a video based industrial safety application. Due to its simple structured processing path, hardware can be designed in a cost effective manner. The robustness of hand detection can be increased by exploiting and spacial coherence consequently, modular and thus scalable versions can be designed which can be adapted to the special requirements of a particular field of application.

References