Project no.: IST-FP6-FET-16276-2
Project full title: Learning to emulate perception action cycles in a driving school scenario
Project Acronym: DRIVSCO

Deliverable no: D3.2
Title of the deliverable: Definition of the event-driven communication scheme

Date of Delivery: 25.06.2009
Organization name of lead contractor for this deliverable: UGR
Author(s): Sara Granados, Sonia Mota, Eduardo Ros (UGR)
Participant(s): UGR
Work package contributing to the deliverable: WP3
Nature: R
Version: 25.06.2009
Total number of pages: 9
Start date of project: 1 Feb. 2006  Duration: 42 months

Project Co-funded by the European Commission

<table>
<thead>
<tr>
<th>Dissemination Level</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PU</td>
<td>Public</td>
</tr>
<tr>
<td>PP</td>
<td>Restricted to other program participants (including the Commission Services)</td>
</tr>
<tr>
<td>RE</td>
<td>Restricted to a group specified by the consortium (including the Commission Services)</td>
</tr>
<tr>
<td>CO</td>
<td>Confidential, only for members of the consortium (including the Commission Services)</td>
</tr>
</tbody>
</table>

Revision Notes: (Revised by UGR and SDU)
Delay justification: This deliverable has suffered a delay of 6 weeks.
After achieving Milestone M5 (“Hardware (low-level vision) and software (middle-level vision) communication scheme feasibility and performance (Month 25)” an internal technical report was generated and the communication protocol was decided. But we have delayed the final specification (D3.2) until it was actually implemented and tested to avoid technical problems such as asynchronous communication by the relevance driven communication channel and other potential problematic issues.
Summary

This deliverable describes the communication scheme that takes full advantage of the sparse maps produced by the condensation module/circuit described in D3.1. The communication scheme includes a grid-based sub-sampled data stream (which is regular and consumes and constant bandwidth) and also a relevance-driven data stream which is based on relevance points (marked by external attention-like modules).

Originally in DRIVSCO only the relevance-based data stream was conceived but during the development of the project we have studied and included the fixed grid-based data stream because it allows for dense maps recovery with enhanced accuracy due to its regularization capabilities at low cost (in terms of bandwidth requirements and condensation processing resources). For the sake of clarity it is worth to repeat that although the communication scheme was originally defined as “event-driven communication scheme” which is also the name of this deliverable, the communication is now hybrid in fact, since it includes a regular data stream (with a fixed bandwidth) and a relevance-based data stream (whose bandwidth is more event-driven as originally conceived).

The condensation modules’ functionality have been described and validated in D3.1, therefore, that Deliverable complements this one D3.2.

Content

Summary.................................................................................................................................... 2
Content...................................................................................................................................... 2

1. Introduction....................................................................................................................... 3
2. Motivation within DRIVSCO............................................................................................... 3
3. Communication protocol................................................................................................... 6
   3.1. Conceptual communication channels ........................................................................ 6
   3.2. Physical communication protocol ........................................................................... 6
   3.3. Back-propagation communication channels .............................................................. 8
4. Discussion .......................................................................................................................... 8
5. Bibliography....................................................................................................................... 9
Definition of the event-driven communication scheme
Authors: Sara Granados, Sonia Mota Fernández, Eduardo Ros
{sgranados, smota, eros}@atc.ugr.es
DRIVSCO Project

1. Introduction

The goal of this deliverable is to describe an efficient communication scheme between low level and middle level vision modules. This will take advantage of the condensation scheme described in D3.1. That condensation scheme sub-sampled the images according to a fixed grid (GR) and an input signal that selected a set of relevant points (RP).

The communication scheme attaches low level processing estimators to these selected points. Therefore, in the communication scheme two different data streams are defined: a) Regular grid-based data stream with data related with a pre-defined grid (this represents a fixed and regular subsampling) and b) Relevant-points-driven data stream with data related with selected points that are salient according to external “attentional” signals (defined by external modules).

As shown in D3.1, using interpolation (from grid-based points and relevant points) we can recover dense low level estimation maps from the sparse representation generated by the condensation module. In fact the recovered dense maps are of higher accuracy than the original ones due to the inherent regularization capability of the condensation scheme. Since the decondensation stage is a smart interpolation operation it inherently performs a kind of regularization. The accuracy of the recovered dense maps is particularly enhanced when the low level estimation maps are obtained with dedicated hardware (as the one being designed in DRIVSCO with restricted computation accuracy and based on fixed point arithmetic). The evaluation of this accuracy enhancement is beyond the scope of this deliverable and will be illustrated with specific results in the last period review.

The defined communication scheme allows for also back-propagation signals and estimations from middle level vision modules into low level vision extraction engines as illustrated in section 3.3.

In section 2, the motivation the communication scheme between low level sparse estimation maps and higher processing stages within DRIVSCO is briefly described. Section 3 describes the communication scheme. Section 4 summarizes the characteristics of the described communication scheme and relates it with results already described in D3.1 and other results obtained in the last period of the project.

2. Motivation within DRIVSCO

As already indicated in D3.1, in the framework of the DRIVSCO project, the condensation stage allows for an efficient representation of low level estimation maps
using a sparse representation. In D3.1 we already described and briefly validated how dense representation maps can be recovered from the sparse maps using conventional interpolation methods. This sparse representation maps provided by the condensation module are used for efficiently communicating data between the specific purpose processing engines implemented on the FPGA (or other acceleration technologies) and the PC (in which middle vision models are implemented as software modules). This condensation module is developed in the framework of WP3 (Integration of Low level and middle level cues) and is integrated with the computing system developed in WP1. The processing engine under development in WP1 produces a dense map of low level vision cues (optical flow, disparity, local energy, phase and orientation). The condensation module described in the deliverable D3.1 condenses this dense map into a sparse representation which is compatible with higher level processing stages towards structured visual events (SVEs) under development in WP4. Furthermore, once the condensation module is integrated on-chip with the low-level vision extraction engines, it provides a sparse representation map that optimizes the bandwidth of the communication channel between on-chip processing and higher processing stages (defined in software). This requires an efficient communication scheme which can be considered partially event-driven since it includes a data stream which is based on externally marked “relevant points”. Initially, the DRIVSCO project considered mainly just this event-driven communication data stream, but during the development of the project we have seen that at low cost (in terms of bandwidth and condensation processing resources) we can also include a fixed grid sub-sampling map which facilitates homogeneous recovering of the original dense maps produced by the low level vision extractors (optical flow, stereo, local contrast descriptors, etc). See D3.1 and Figure 1 for illustrative results.
In D3.1, we have already described that the condensation scheme includes a regular sub-sampling method (for instance using a grid of 3x3 or 5x5). But some points of the image may have specific relevance or their “vision cues” (motion, stereo, etc) may be of significant confidence. This is hard to estimate a priori, but for validation purposes in D3.1 we have included a simple “edge extraction” model [1], which can also be replaced or complemented with an intrinsic dimension estimator [2, 3, 4] or other selection driving signals. We can use local energy as “relevant information indicator” which is very much related with the confidence measurements of the stereo and optical flow estimations of the vision front-end. Other relevance estimators (such as attention driving signals as Time-To-Contact (TTC) estimations, IMOs, etc) can also be used.

The background sub-sampling technique forces the model to take samples even in smooth surfaces allowing for interpolation of the condensed representation to recover the original images. This sub-sampling data is called “grid-based data” while the samples taken at points with high relevance (such as high “local energy” or other “relevance indicators”) will be called “event-driven data or relevance-driven”.

In D3.1 we have already illustrated how with a simple local-energy-based relevant points scheme and a regular grid of 5x5 the communication bandwidth is reduced to a 7% of the original bandwidth requirements and the accuracy of the recovered maps (with simple interpolation methods) is maintained high. Furthermore, as indicated above, the accuracy of the recovered maps is even higher than the original ones (due to the inherent regularization capability of the condensation scheme).
3. Communication protocol

3.1. Conceptual communication channels

Once condensed, the information must be sent to higher processing stages using an efficient protocol. We have used a hybrid “regular and event driven” protocol. Thus, we have developed a communication protocol which sends information through two different channels:

- Grid channel: regular information related to the fixed grid is sent by this channel. Therefore low level vision estimations (such as optical-flow) from points of this grid are sent. This represents a regular data stream. The bandwidth requirements are low (due to the sub-sampling applied) and constant.

- Event-driven channel: following an AER (Address Event Representation) protocol, relevant points indicated by the “relevance-driven selectors” are marked. Low level vision estimators from these points are sent (attached to the source addresses). This represents an event-driven data stream. The bandwidth is defined by the number of relevant points driven by external “attention-like” signals.

In order to fulfill Milestone M5: “Hardware (low-level vision) and software (middle-level vision) communication scheme feasibility and performance (Month 25)” we studied different schemes for doing the physical communication. Sending a binary matrix for the relevant points indicating if it is active or not (considered relevant or not by the external module) or sending directly the address of each relevant point (AER-like scheme). We obtained that for images of 1024x1024 if the relevant points were less than 4% of the whole image the bandwidth consumed by the AER scheme was below the one required by a mask-based approach (in which the binary mask needs to be transmitted in each image). Therefore we adopted the AER-like approach for the relevant points.

3.2. Physical communication protocol

The communication between the low level modules and the middle ones uses the PCIe port which is a serial communication channel. This means that the communication must be done over shared memory as shown in Figure 2. This is generalizable to other physical communication technologies.
On the one hand, the amount of regular information related with the fixed grid is constant, depending just on the image size. The grid is fixed; therefore we avoid storing explicitly the addresses of the different pixels of the grid. The software driver that takes the data will also have a replicated grid-definition look-up-table (in this way, it can locate each estimation in its corresponding position). Therefore, in the shared memory we store only the grid values as a list of estimations, i.e., a list of data, avoiding sending explicit addresses (which would be repetitive for each image and waste communication bandwidth). This option reduces the bandwidth requirement significantly. It only requires pre-defining the grid on the condensation module and on the software receiver driver.

On the other hand, the relevant points are variable and they are stored using an AER-like scheme, i.e., one register per datum is used containing explicitly its address (row and column number) and its values (attached low level vision estimations).

The final physical implementation is shown in Figure 3. As we can see the regular information has a limited storage space and the relevant point’s space is only restricted by the SRAM size. The figure also provides an example of memory use for a 512x512 image assuming 1% of relevant points.

<table>
<thead>
<tr>
<th>SRAM</th>
<th>(Words)</th>
<th>KB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular information (Grid Points)</td>
<td>0x00000000 0x00000DA7 0x00000DA8</td>
<td>15.8KB</td>
</tr>
<tr>
<td>Relevant Points</td>
<td>0x00002C5F</td>
<td>35.4KB</td>
</tr>
</tbody>
</table>
3.3. Back-propagation communication channels

A similar communication scheme is used to back-propagate middle level estimations into low-level representation maps. The physical communication in this case is only done adopting an AER scheme (Address-Event-Representation). Two different channels are defined:

a. Relevance signals. This drives the relevance-based data. This makes possible back-propagating signals of interest (such as Time-To-Contact, or IMOs) and uses them to select areas or points of interest in which low level estimations are desired.

b. Middle-level estimations such as highly confident estimations. This data can be directly used at the condensation module by weighted averaging it locally with the grid-based data. This modifies the low level grid-based estimations.

The physical communication scheme is done in the same manner as feed-forward communication, i.e. through shared memory.

4. Discussion

The communication scheme defined in this Deliverable takes full advantage of the sparse representation maps produced by the condensation module/circuit developed in WP3. It optimizes the communication bandwidth by transferring only the output of the condensation module (sparse representation) instead of the original low-level vision dense maps of extracted cues.

The communication protocol is flexible enough to allow for a regular grid-based data stream and a “relevance-driven” (event-driven) data stream co-exist in the framework of the low-level to middle level communication. Furthermore, it also allows for the backpropagation of signals and high confident estimations from middle-level modules into low level vision representation maps.

In D3.1 we have illustrated the reduction bandwidth requirements facilitated by the condensation module and actually used by the communication scheme described in this deliverable (D3.2). D3.1 also illustrated how the low-level dense maps can be recovered from the sparse maps provided by the condensation module. But in the third period of the project we have even gone beyond this. We have also studied and demonstrated that the accuracy of the recovered maps (from the sparse representation) is even higher than the original (dense low level maps). This is supported by the regularization capability of the condensation scheme (specific results will be shown in the last period review).
Therefore the condensation representation provides different advantages:

a. Optimization of the communication bandwidth used by the here described communication protocol.

b. Enhancement of the accuracy of the transferred maps due to the inherent regularization capability of the condensation scheme. This is a feature not expected at the beginning of the project but facilitated by the inclusion of the hybrid representation scheme (regular grid-based and sparse relevance-based estimators).

c. Facilitation of back-propagation scheme to use middle level cues and fuse them into low level estimation maps. It also allows for back-propagating attention-like signals to drive the relevance-based data stream.

5. Bibliography


