# A GROUP THEORETICAL TOOLBOX FOR COLOR IMAGE OPERATORS

*Reiner Lenz, Thanh Hai Bui*

Department Science and Technology
Linköping University, SWEDEN

*Koichi Takase*

Graduate School of Science and Technology
Chiba University, JAPAN

**ABSTRACT**

In this paper we describe how to use the direct product of the dihedral group $D(4)$ and the symmetric group $S(3)$ to automatically derive low-level image processing filter systems for RGB images. For important classes of stochastic processes it can be shown that the resulting operators lead to a block-diagonalization of the correlation matrix.

We will show that the group theoretical derivation of the operators leads to a very fast implementation consisting mainly of additions and subtractions. They can therefore be implemented in fast graphics computation hardware such as a GPU. We then illustrate the block-diagonalization property in an experiment where we used 20,000 subimage patches collected from 10,000 random images in a large image database. The very short execution times make these operators suitable for applications where many images have to be processed. Typical examples are video processing and content-based image database search. We describe one example where we use the operators to compute content based descriptors of images. These descriptors are currently used in one search mode in an image database browser operating on a database with more than 750,000 images. The group theoretical tools used to derive these filters are very general and can directly be applied for other types of image data. Examples are the following generalizations of the methodology: filter systems for multiband images with more than the ordinary three RGB-channels or images with other grid geometries such as hexagonal sampling.

## 1. INTRODUCTION

Group theoretically defined transforms are important tools in many low-level digital signal processing applications. The most widely used examples are the DFT and the FFT but also wavelet transforms can be understood in terms of group theory. These transforms are attractive for many reasons, among them the possibility to derive the transforms algorithmically and efficient implementations that make them suitable for tasks that require very fast analysis of very large amounts of data.

In this paper we describe how the general tools of group representation theory can be used to design filter systems that analyze multichannel images. We propose an algorithm that makes it possible to use numerical and symbolical math-programs such as Matlab and Maple to derive a simplified version of these filter systems. These simplified systems are equivalent to the original filter systems but they consist only of additions and subtractions. The resulting filter systems have an intrinsic structure that allows a fast, hierarchical implementation similar to the FFT. We illustrate the application of these filter systems in the framework of color based image database search and we will discuss their connection to the Principal Component Analysis of stochastic processes with group theoretical symmetries.

## 2. GROUP THEORETICAL FILTERS

### 2.1. Dihedral and permutation groups, representations

Consider a small patch of an image. The structure of this patch specifies the local pattern of the image at the patch location. It is reasonable to assume that, for a large collection of patches, the rotated or flipped version of the same patch will appear with the same probability (see [1, 2] for an early application of these symmetries). A similar assumption can also be made with the permutations of color channels within a patch. We will thus assume that all transformations that map a collection of patches into itself will produce patterns of approximately the same probability. For the conventional RGB images on a square grid the rotation and flip transformations are described by the dihedral group $D(4)$, while the color channel permutations are described by the permutation group $S(3)$. The dihedral group $D(4)$ consists of all rotations with angle $\{90 \cdot k : 0 \leq k \leq 3\}$ degrees and all reflections around the symmetry axes of a square. It has eight elements. The group of all permutations

The financial support from Vinnova (the Swedish Agency for Innovation Systems), the Swedish Research Council and STINT (The Swedish Council for International Cooperation in Research and Higher Education) is gratefully acknowledged. The commercial version of the image database used in the paper is owned by Matton AB, Stockholm, Sweden and available under http://www.matton.se

0-7803-9134-9/05/$20.00 ©2005 IEEE

of three elements is $S(3)$ has six elements, we denote them by $\pi_1, \dots, \pi_6$. The symmetry group we will use in the following is $G = D(4) \times S(3)$ consisting of all pairs $(g, \pi)$ where $g \in D(4)$ and $\pi \in S(3)$. It will be used to describe all possible spatial/color transformations of a patch.

Next take N points on a square grid that are invariant under the dihedral group and on each point an RGB vector. The signal vector has thus dimension 3N and is of the form

$$\boldsymbol{s} = (\boldsymbol{R_1}, \dots, \boldsymbol{R_N}, \boldsymbol{G_1}, \dots, \boldsymbol{G_N}, \boldsymbol{B_1}, \dots, \boldsymbol{B_N})$$

where $(\boldsymbol{R_k}, \boldsymbol{G_k}, \boldsymbol{B_k})$ is the RGB vector at position $k$. Applying a dihedral transformation to the locations and a permutation to the RGB channels produces a new signal of the same type. Each element $(g, \pi)$ defines thus a transformation $\boldsymbol{T}(g, \pi)$ that maps a 3N-dimensional signal $s$ to another 3N-dimensional signal $\boldsymbol{T}(g, \pi)s$. These transformations preserve the group operation in the sense that

$$\boldsymbol{T}(g_k, \pi_k)\boldsymbol{T}(g_l, \pi_l) = \boldsymbol{T}(g_k g_l, \pi_k \pi_l)$$

for all $(g_k, \pi_k), (g_l, \pi_l) \in G$. Such a construction is called a representation of a group. Within the 3N-dimensional vector space there might exist subspaces that are also invariant under all such transformations $\boldsymbol{T}(g, \pi)$ and operating on such a subspace we obtain representations of the group on the smaller space. The smallest invariant spaces under a group defines the so called irreducible representations of the group. These irreducible representations are known for the most important groups and methods to construct them can be found in the literature (see [3–5] or any other book on group representations).

The irreducible representations of $G = D(4) \times S(3)$ are the Kronecker products of the irreducible representations of $D(4)$ and $S(3)$. It can be shown that there are five irreducible representations for $D(4)$ (four operating on a one-dimensional and one operating on a two-dimensional space). There are three irreducible representations for $S(3)$ (two of dimension one and one of dimension two). There are thus 15 irreducible representation of G and they have dimensions one, two or four. We denote them by $\{\boldsymbol{T}_k^G\}_{k=1}^{15}$ and define

$$\boldsymbol{T}^G_{3(\nu-1)+\mu} = \boldsymbol{T}^D_\nu \times \boldsymbol{T}^S_\mu \tag{1}$$

### 2.2. Derivation of the filters

The strategy behind group theoretical filter design is to construct a coordinate system in the signal space that is well-adapted to the transformation group under consideration. This means that the signal space is split into subspaces such that each subspace is invariant under all the transformations in the group and such that the transformation on each subspace is as simple as possible. Filter values will thus transform like the irreducible representations. Instead of constructing filter functions that are defined on the 3N-dimensional signal space we simplify the construction by first dividing the original 3N-dimensional space into smaller invariant subspaces. A special type of invariant subspaces are the orbits. These orbits are obtained by applying all dihedral transformations to a given grid point and then applying all permutations to all channels on a given point. It can be shown that there are four types of orbits that can be generated by the dihedral group (see Fig. 1):

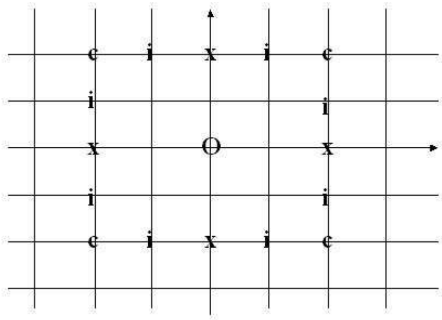


**Fig. 1**. Different orbit types for the dihedral group

- **origin (o)**: Consisting of the single point at the origin
- **corner orbits (c)**: These are all points on the corner of squares centered by the origin. There are four points on such an orbit.
- **axis orbits (x)**: Containing 4 points on the coordinate axes
- **inner orbits (i)**: These are the remaining orbits and have 8 elements.

The signals defined on these orbits define invariant subspaces but these subspaces can in general be subdivided further since we saw from the general theory that the smallest invariant subspaces of the group $D(4)$ have dimensions one and two. The orbit spaces have dimensions 3, 12 or 24 and in a first stage we consider them separately. The transformations of $S(3)$ on the color channels are applied to a single color pixel and can thus be separated from the transformations of the dihedral group $D(4)$. The construction of the filter functions can either be done "manually", i.e. by using the well-known irreducible representations of the dihedral and the permutation group or they can be obtained algorithmically by using procedures from the theory of group representations (see [5] for a description). In that case programs like Maple and Matlab can be used to construct the filter systems for any given invariant signal space. The result shows that the filters operate as follows:

1. Compute from the input RGB image three combined images: R+G+B (originating in the trivial representation), R-B, and G-B (from the 2-D representation).

2. Each of these 3 combinations is then filtered with the spatial filter defined by the operation of the dihedral group $D(4)$. This results in 1, 4, or 8 filter images (for each of these combinations) for the case of origin, corner, or inner orbits respectively.

3. The corner orbit acts on 4 pixels, and produces 4 filter images for each combined image given by step 1. They are computed by the following functions:

$$\begin{aligned} p_1 + p_2 + p_3 + p_4 \\ p_1 - p_2 + p_3 - p_4 \\ p_1 - p_3 \\ p_2 - p_4 \end{aligned} \tag{2}$$

   where $p_i$ denotes the pixel value of each computed combined image, with index $i$ increases clockwise.

4. For the inner orbit consisting of 8 pixels we get filter systems that are of the same simple form as the filters in Eq. (2).

Taking into account the ordering of the pixels we see that in Eq. (2): the first filter (1,1,1,1) computes the sum over the pixel values. The second filter (1,-1,1,-1) is a diagonal line detection filter. If we combine the last two by addition and subtraction we find that they generate the filters (1,1,-1,-1) and (1,-1,-1,1) corresponding to y- and x-gradient filters. Note that the simplified filters are not orthonormal. To obtain the filter results corresponding to the unitary representations of the group the filter vectors have to be corrected by a matrix multiplication.

## 3. IMPLEMENTATION AND EXPERIMENTAL RESULTS

It is known that group-theoretically constructed filter systems permit fast implementations. The most common example is the fast FFT implementation of the DFT (see [6, 7] for a general discussion). A fast implementation of the filter systems described above can be obtained by reusing the intermediate computations with the spatial operators described in Eq. (2). For example, the spatial filter of the corner orbit can be computed as follows:

- The four spatial combinations can be computed using the following tree structure: $q_1 = p_1 + p_3; q_2 = p_2 + p_4; s_1 = q_1 + q_2, s_2 = q_1 - q_2, s_3 = p_1 - p_3, s_4 = p_2 - p_4$ where $p_k$ denotes the k-th pixel in the sequence, $q_l$ denotes an intermediate result and $s_m$ is the m-th final filter result for that channel.

- We need 6 additions/subtractions for one channel to compute the basic filter results and 16 additions/subtractions for the combination of the final results. The full implementation consists thus of 34 additions/subtractions for all twelve filter functions.

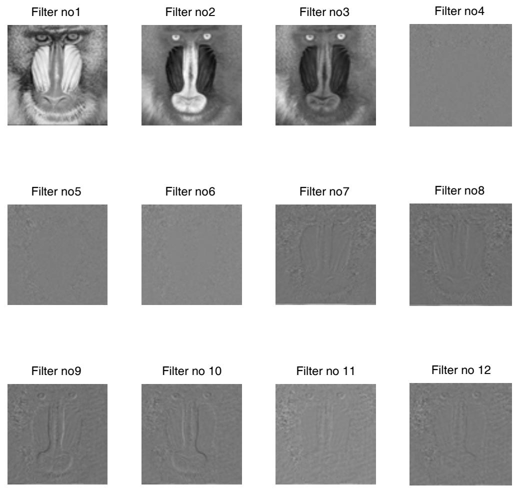



**Fig. 2**. Corner-orbit filters applied to the Baboon image

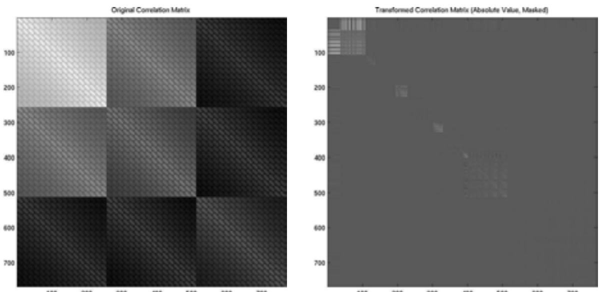
This highly parallelizable, linear structure of the algorithm allows us to implement the filters in special hardware such as a GPU. The computation time when using Geforce6600 GT GPU was less than 0.0056 second (180FPS) for computing all twelve filter results for an image of size $512^2$. An example of applying the twelve filters defined on the corner orbit to the Baboon image[1] is shown in Fig. 2 (the raw filter result shifted and scaled by the Matlab command *imagesc*). An interpretation of some of these filters is the following: No 1,2,3: Smoothing in the gray value, the R-B and G-B images, No 9: diagonal spatial gradient in the R-B image.

Filter systems defined on different orbits can be combined to generate larger filters. How filter originating in the same irreducible representation but computed from different orbits should be combined cannot be answered by group theoretical arguments alone.

If the assumption that all transformations are equally likely is correct, then it can be shown that the group theoretically derived filters will block-diagonalize the correlation matrix. To check if this assumption holds for realistic image sources we made the following test: From a large database with more than 750,000 images we select 10,000 random images[2]. From each image we collect randomly 2 patches of size $16 \times 16$ pixels. The signal vector had thus dimension $16 \times 16 \times 3 = 768$. Collecting the group theoretical filters in the matrix $\boldsymbol{M}$) we compute the transformed correlation matrix $\boldsymbol{M'CM}$. Fig. 3 shows the original correlation matrix and the absolute value of the transformed correlation

[1] see `http://rsb.info.nih.gov/ij/images/baboon.jpg`

[2] a browser interface to the database is available at `http://www.itn.liu.se/media/vinnova/cse.php`



(a) Original correlation matrix $\boldsymbol{C}$ (b) Masked block-diagonal correlation matrix $\boldsymbol{M'CM}$

**Fig. 3**. Original correlation matrix and its masked block diagonalized version

matrix where we have masked out the (dominating) upper left corner to visualize the remaining entries.

## 4. AN APPLICATION IN IMAGE DATABASE RETRIEVAL

Content based image retrieval from large image databases requires fast processing methods like the filter systems described here. One of the search modes in our search engine[3] uses these filter systems as follows: The filters operate on the innermost corner and inner orbits resulting in a 36-dimensional signal vector providing information about the local structure of the image. They define a $36 \times 36$ correlation matrix with a block-diagonal structure defined by the group. We use only the blocks in the diagonal of this matrix to produce a descriptor vector of length 130 for each image (The diagonal block set of this matrix consists of one $8 \times 8$ block, two $4 \times 4$ blocks, seven $2 \times 2$ blocks, and six single values). The symmetry property will therefore reduce the length of the descriptor vector from 130 to 83. We apply the PCA then to compress these descriptors. These PCA-compressed descriptors are used for the retrieval. An example of a search result given by our CBIR system using this search mode is illustrated in Fig. 4.

## 5. CONCLUSION

We first described how the general tools of group theory can be used to automatically derive the image filters. This lead to very fast filter systems that block-diagonalize the correlation matrix of the filtered signal vectors. This partial PCA and the very fast execution times make them suitable for applications where speed is critical such as real time filtering or image database processing. Here we described the filter systems for RGB images defined on square grids. This is only a demonstration of the general framework. Multichannel images, or images defined on other types of regular grids can also be investigated with the same group theoretical tools (now with symmetry group $D(n) \times S(m)$ where $n$ defines the grid symmetry and $m$ the number of channels).

[3] `http://www.itn.liu.se/media/vinnova/cse.php`

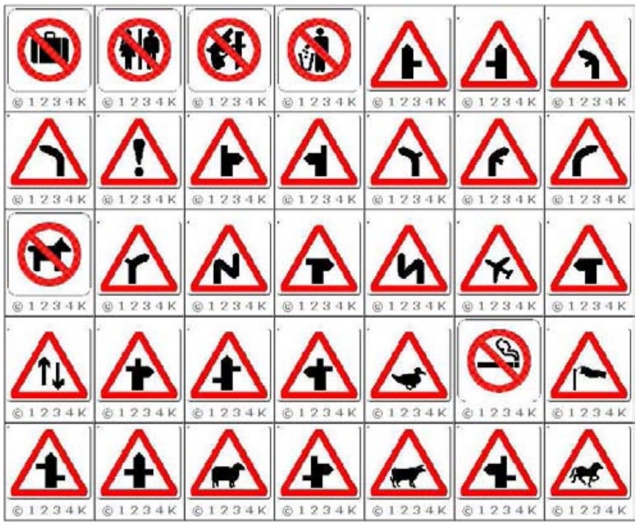


**Fig. 4**. An example of image search results using group filter correlation matrix as content based descriptor. Search query URL: http://www.itn.liu.se/media/vinnova/cse.php?Server=V5 &ImageIndex=308150

## 6. REFERENCES


[1] R. Lenz, "Using representations of the dihedral groups in the design of early vision filters," in *ICASSP 1993*, pp. V165–V168.

[2] ——, "Investigation of receptive fields using representations of dihedral groups," *J. Vis. Comm. Im. Rep.*, vol. 6, pp. 209–227, 1995.

[3] J. P. Serre, *Linear Representations of Finite Groups*. Springer, 1977.

[4] W. Fulton and J. Harris, *Representation Theory*. Springer, 1991.

[5] A. Fässler and E. L. Stiefel, *Group theoretical methods and their applications*. Birkhäuser, 1992.

[6] D. N. Rockmore, "The FFT: An algorithm the whole family can use," *Comp. Sci. & Eng.*, vol. 2, pp. 60–64, 2000.

[7] D. N. Rockmore and D. M. Healy, Jr, Eds., *Modern Signal Processing*. Cambridge U. Press, 2004.