AFFINE MORPHOLOGY AND AFFINE SIGNAL MODELS

Petros Maragos

Division of Applied Sciences, Harvard University, Cambridge, MA 02138, USA.

ABSTRACT: Affine signal transformations are useful for modeling self-similar structures in fractal images and shape deformations in visual motion. In the first part of this paper a theoretical framework, called affine morphology, is developed to analyze parallel and serial superpositions of affine image transformations. Affine morphology unifies and extends translation-invariant morphological image transformations and their rotation/scaling-invariant generalizations by using action of affine groups on lattices. Several theoretical aspects of affine morphology are explored for binary images. In the second part of the paper, the affine transformations are extended to gray-level images and arbitrary signals, and affine models are developed by using a sum superposition of affine signal transformations. A solution is then given to the problem of estimating the parameters of this sum-affine model using least squares algorithms, and some applications are outlined for image and speech signal processing.

1 Introduction

The analysis of useful classes of image transformations and the design of systems based on them is of central importance in image processing and computer vision. One such traditional class has been the class of linear and translation-invariant image transformations, with applications mainly to image filtering, restoration, and coding. Another class of growing popularity is the collection of all morphological image operations, which are nonlinear and translation-invariant; we call this whole area *translational morphology*. Morphological operations have found many applications in nonlinear image filtering, image enhancement, feature detection, shape analysis, and pattern recognition; see [12] for a recent survey. However, not all image processing tasks are translation-invariant. There are cases where invariance under rotation or scaling is needed. Recently, morphological operations for binary images have been introduced that are invariant under rotation and scalar multiplication; we call this area *polar morphology*. However, while both translational and polar morphology have shortcomings because they lack each other's invariance characteristics, neither of them can represent a broad class of image transformations, that occur in visual motion and are used in fractal image modeling. These are the *affine* image transformations.

Therefore, in this paper we focus on the class of affine image transformations and develop a theoretical framework that can (i) unify translational and polar morphology as special cases of what we call *affine morphology*, and (ii) provide analytic tools for a variety of shape deformations used to model fractal images or interpret visual motion. In these applications we have superpositions of translated, scaled and rotated versions of the original image. Whereas translational morphology can model only translations and polar morphology can model only uniform (i.e., same with respect to both axes) scaling/rotation, affine morphology can simultaneously model both translations and arbitrary rotations with arbitrary scalings. The first part (Sections 2 and 3) of the paper explores (i) and (ii) for binary images. (iii) In the second part (Section 4) of the paper, in an effort to extend the above ideas to gray-level images, we develop a *sum-affine* model that is capable of modeling a signal as an additive superposition of affine domain-transformed and affine amplitude-transformed versions of another signal (in which case the model generalizes convolutions) or of itself (in which case the model is *self-affine*). Then we provide least-squares algorithms to find the parameters of this model and we outline its potential application to modeling self-similarities in fractal images and other signals, e.g., speech.

In the rest of this introduction we outline some background ideas on affine image transformations and their applications, and define the basic operations of translation and polar morphology.

AFFINE IMAGE TRANSFORMATIONS: Consider the class of binary images, represented by subsets of the two-dimensional (in short, 2-dim) Euclidean space $\mathbf{R}^2$. Then a broad and useful class of image transformations can be constructed based on affine point mappings $w(x) = Mx + t$, $x \in \mathbf{R}^2$, where $t \in \mathbf{R}^2$ and $M \in \mathbf{R}^{2\times 2}$ is a 2×2 real matrix. The vector-valued map w can be extended to the set $\mathcal{P}(\mathbf{R}^2)$ of subsets of $\mathbf{R}^2$. Thus by an *affine image set transformation* we mean a set-valued mapping

$$w(X) = \{Mx + t : x \in X\} \quad , \quad X \subseteq \mathbf{R}^2. \tag{1}$$

Affine transformations have been (explicitly or implicitly) used often for modeling 2-dim or 3-dim shape deformations with applications to computer vision, photogrammetry, and mathematical psychology. For example, the motion of a rigid body in 3-dim space can be modeled by a 3-dim affine transformation, which when projected down to the 2-dim camera's image plane or the eye's retina, induces (under certain restrictions on the type of motion and projection) 2-dim affine transformations between the projected image sets [1, 4]. Several aspects of visual perception (e.g., shape and size constancy, motion) have been modeled by Lie groups whose elementary finite transformations are 2-dim affine mappings [8]. There has also been considerable interest and research work in inverting an affine transformation either for relating two coordinate systems [9], or for recovering motion parameters [4, 11], or for recognizing a transformed shape by reffering it to an undeformed original shape [18, 11]. Note that in the above mentioned applications the overall image transformation was modeled by a single affine transformation or in visual motion by a composition (and thus a *serial* superposition) of several affine transformations. However, in modeling fractal images [10, 3], a union, i.e., a *parallel* superposition of affine transformations has been used to model fractal images as the attractors of chaotic dynamical systems.

TRANSLATIONAL MORPHOLOGY: All binary image transformations of traditional mathematical morphology [13, 15] are translation-invariant and stem from the dilation $\oplus$ and erosion $\ominus$, which are, respectively, the union and intersection of translations $X \pm b = \{x \pm b : x \in X\}$ of the image set $X \subseteq \mathbf{R}^2$ by vector points $b = (b_1, b_2)$ of the structuring element $B \subseteq \mathbf{R}^2$:

$$X \oplus_T B \;=\; \bigcup_{b\in B} X + b = \bigcup_{(b_1,b_2)\in B} \{(x_1 + b_1, x_2 + b_2) : (x_1, x_2) \in X\} \tag{2}$$

$$X \ominus_T B \;=\; \bigcap_{b\in B} X - b = \bigcap_{(b_1,b_2)\in B} \{(x_1 - b_1, x_2 - b_2) : (x_1, x_2) \in X\} \tag{3}$$

Note that we have added to the usual notation of dilation $\oplus$ and erosion $\ominus$ a subscript T to emphasize the fact that both the classical dilation and erosion are obtained from translations of planar points. We shall call them *translational dilation and erosion*, and they are identical to classical Minkowski set addition and subtraction defined in [5]. This paper also examines other forms of dilation and erosion for which we will use the symbols $\oplus$ and $\ominus$ with different subscripts. Two very important properties that the translational dilation and erosion satisfy are: 1) translation-invariance and 2) distributivity over union or intersection, respectively: i.e.,

$$(X + t) \oplus_T B = (X \oplus_T B) + t \quad ; \quad (\bigcup_{j\in J} X_j) \oplus_T B = \bigcup_{j\in J} X_j \oplus_T B \tag{4}$$

$$(X + t) \ominus_T B = (X \ominus_T B) + t \quad ; \quad (\bigcap_{j\in J} X_j) \ominus_T B = \bigcap_{j\in J} X_j \ominus_T B \tag{5}$$

where J is some index set, countable or not.

POLAR MORPHOLOGY: To create rotation- and scaling-invariant mophological operations, Heijmans and Roerdink [6, 14] developed the *polar morphology* by observing that translational erosion and dilation depend only on vector addition, which forms a commutative group in $\mathbf{R}^2$ and induces translations of planar sets by vectors. They replaced this translation group structure in $\mathbf{R}^2$ with the "*polar group*", which is the set of nonzero planar points represented in polar coordinates (r, θ) and equipped with the group operation $+_P$ defined by $(r_1, \theta_1) +_P (r_2, \theta_2) = (r_1 \cdot r_2, \theta_1 + \theta_2)$; i.e., in the polar group two planar points are combined by multiplying their magnitudes and adding (modulo 2π) their angles. Then they defined the *polar dilation and erosion* as:

$$X \oplus_P B = \bigcup_{(\rho,\phi)\in B} X +_P (\rho, \phi) = \bigcup_{(\rho,\phi)\in B} \{(r \cdot \rho, \theta + \phi) : (r, \theta) \in X\} \tag{6}$$

$$X \ominus_P B = \bigcap_{(\rho,\phi)\in B} X +_P (\rho^{-1}, -\phi) = \bigcap_{(\rho,\phi)\in B} \{(r/\rho, \theta - \phi) : (r, \theta) \in X\} \tag{7}$$

The polar dilation and erosion are: 1) invariant under rotation and/or scalar multiplication and 2) distributive with respect to union or intersection, respectively: i.e.,

$$(X +_P (\rho, \phi)) \oplus_P B = (X \oplus_P B) +_P (\rho, \phi) \quad ; \quad (\bigcup_{j\in J} X_j) \oplus_P B = \bigcup_{j\in J} X_j \oplus_P B \tag{8}$$

$$(X +_P (\rho, \phi)) \ominus_P B = (X \ominus_P B) +_P (\rho, \phi) \quad ; \quad (\bigcap_{j\in J} X_j) \ominus_P B = \bigcap_{j\in J} X_j \ominus_P B \tag{9}$$

Polar morphological operations are useful in problems without translational symmetry where the image operations need to be invariant under rotations and uniform scaling. An example is images photographed with a fish-eye (wide-angle) lens; there we need rotation/scaling-invariance with respect to the optical center of the image (intersection of camera axis and image plane) instead of translation-invariance.

2 Group Action Morphology on Lattices

For generalization purposes, an abstract viewpoint is adopted in this section. Let $S \subseteq \mathbf{R}^2$. The set $\mathcal{P}(S) = \{X : X \subseteq S\}$ of all (binary images) subsets of S is viewed as an atomic Boolean complete lattice, where the supremum and infimum lattice operation are the set union and intersection, respectively. The *atoms* of $\mathcal{P}(S)$ are the singletons $\{x\}$, $x \in S$. Image transformations from $\mathcal{P}(S)$ to $\mathcal{P}(S)$ are called (set) *operators*, and $\mathcal{O}$ denotes the set of all such operators.

2.1 Preliminaries from Morphology on Lattices

Let $\alpha, \beta \in \mathcal{O}$. Their *composition* is the operator $\alpha\beta(X) = \alpha(\beta(X))$, $X \subseteq S$. Set $\subseteq$ induces a similar ordering between operators, i.e., $\alpha \subseteq \beta$ means $\alpha(X) \subseteq \beta(X)$ for all $X \subseteq S$. This ordering makes $\mathcal{O}$ a complete lattice. α^{-1} denotes the inverse, if it exists, of the operator α. Let $id(X) = X$ denote the *identity* operator on $\mathcal{P}(S)$.

DEFINITION 1 . (Serra [16]) Let $\beta \in \mathcal{O}$. Then:
(i) β is *increasing* if $X \subseteq Y \implies \beta(X) \subseteq \beta(Y)$ for all $X, Y \subseteq S$.
(ii) β is a *dilation* if $\beta(\bigcup_{X\in\mathcal{K}} X) = \bigcup_{X\in\mathcal{K}} \beta(X)$ for all $\mathcal{K} \subseteq \mathcal{P}(S)$, i.e., if β distributes over $\cup$.
(iii) β is an *erosion* if $\beta(\bigcap_{X\in\mathcal{K}} X) = \bigcap_{X\in\mathcal{K}} \beta(X)$ for all $\mathcal{K} \subseteq \mathcal{P}(S)$, i.e., if β distributes over $\cap$.
(iv) β is an *opening* if it is increasing, idempotent (i.e., $\beta\beta = \beta$) and anti-extensive (i.e., $\beta \subseteq id$).
(v) β is a *closing* if it is increasing, idempotent and extensive (i.e., $\beta \supseteq id$).

An operator $\alpha \in \mathcal{O}$ is an *automorphism* on $\mathcal{P}(S)$ if it is a bijection such that both α and α^{-1} are increasing. Note that dilations and erosions are increasing. The following can be shown about automorphisms:

PROPOSITION 1 . *Let $\alpha \in \mathcal{O}$. The following four statements are equivalent:*
(i) α is an automorphism. (ii) α is a bijection, dilation and erosion.
(iii) α is a bijection and dilation. (iv) α is a bijection and erosion.

Let $\delta, \varepsilon \in \mathcal{O}$. Then the pair (ε, δ) is called an *adjunction* [7] (or a *morphological duality* [16]) if $\delta(X) \subseteq Y \iff X \subseteq \varepsilon(Y)$ for all $X, Y \subseteq S$. Given a dilation δ, there is a unique erosion ε such that (ε, δ) is adjunction, and vice-versa. The following summarizes some useful facts about adjunctions:

PROPOSITION 2 . (Serra [16], Heijmans & Ronse [7]) *Let (ε, δ) be an adjunction. Then:*
(i) δ is a dilation and ε is an erosion. (ii) $\delta\varepsilon\delta = \delta$ and $\varepsilon\delta\varepsilon = \varepsilon$. (iii) $\varepsilon\delta \supseteq id$ and $\delta\varepsilon \subseteq id$.
(iv) If $(\varepsilon_j, \delta_j)$, $j \in J$, are adjunctions, then $(\bigcap_j \varepsilon_j, \bigcup_j \delta_j)$ is an adjunction.
(v) If α is an automorphism, then (α, α^{-1}) is an adjunction.

2.2 Group Action Morphology

We unify and extend the translational and polar morphology by using action of groups on the set $\mathcal{P}(S)$. Let $(G, \star)$ be a group with group operation $\star$, identity element e, and the inverse of any $g \in G$ denoted as g^{-1}.

DEFINITION 2 . An *action* of the group G on the set $\mathcal{P}(S)$ is defined as any mapping $\Delta : \mathcal{P}(S) \times G \to \mathcal{P}(S)$ such that, if we write $\Delta_g(X)$ for $\Delta(X, g)$,
A1. $\Delta_h(\Delta_g(X)) = \Delta_{h\star g}(X)$, for all $g, h \in G$ and all $X \subseteq S$.
A2. $\Delta_e(X) = X$, for all $X \subseteq S$.

Thus given an action Δ, each group element $g \in G$ induces an image set transformation $\Delta_g : \mathcal{P}(S) \to \mathcal{P}(S)$, and thus each Δ_g is an operator in $\mathcal{O}$. The following can easily be shown about Δ:

PROPOSITION 3 . *(i) The operator Δ_g is a* bijection *for all $g \in G$.*
(ii) Δ_g is an automorphism if and only if it is a dilation or an erosion.
(iii) Δ_g is an automorphism if and only if $(\Delta_g, \Delta_{g^{-1}})$ is an adjunction.

The set $\mathcal{G} = \{\Delta_g : g \in G\}$ is a group under composition $\Delta_h \Delta_g = \Delta_{h\star g}$. It is a group of bijections on $\mathcal{P}(S)$. The identity of the group is $\Delta_e = id$ and the inverse of Δ_g is $(\Delta_g)^{-1} = \Delta_{g^{-1}}$. If G is commutative, then $\mathcal{G}$ is a commutative group too. Further, if we define the mapping $\Xi : G \to \mathcal{G}$ by $\Xi(g) = \Delta_g$, then Ξ is a *group homomorphism.* The kernel of Ξ is $Ker(\Xi) = \{g \in G : \Delta_g = id\}$. As is well known, the group $\mathcal{G}$ is isomorphic to the quotient group $G/Ker(\Xi)$. If $Ker(\Xi) = \{e\}$, then Ξ is a bijection, $\mathcal{G}$ is isomorphic to G and the action Δ is called *faithful.*

Based on the group action, we define two set operations $\oplus_\Delta : \mathcal{P}(S) \times \mathcal{P}(G) \to \mathcal{P}(S)$ (called the *group-action Minkowski set addition*) and $\ominus_\Delta : \mathcal{P}(S) \times \mathcal{P}(G) \to \mathcal{P}(S)$ (called the *group-action Minkowski set subtraction*) as

$$X \oplus_\Delta B = \bigcup_{g \in B} \Delta_g(X) \quad , \quad X \ominus_\Delta B = \bigcap_{g \in B} \Delta_{g^{-1}}(X). \tag{10}$$

$B \subseteq G$ is a set of group elements and plays the role of a "structuring element".

EXAMPLE 1 . Let $S = G = \mathbf{R}^2$ and let '$\star$' be the vector addition. By defining the action as the translation $\Delta_g(X) = X + g$ of $X \subseteq \mathbf{R}^2$ by $g \in \mathbf{R}^2$, we obtain the classical translational morphology because the $\oplus_\Delta$ and $\ominus_\Delta$ operations become identical to the translational dilation and erosion. Note that this action is faithful, and G is isomorphic to the translation group on $\mathbf{R}^2$.

EXAMPLE 2 . Let $S = \mathbf{R}^2 \setminus 0$ and let $(G, \star) = (S, +_P)$ be the "polar group" of Section 1. Let us define the group action by $\Delta_{(\rho,\phi)}(X) = X +_P (\rho, \phi) = \{(r\rho, \theta + \phi) : (r, \theta) \in X\}$, $X \subseteq S$, which rotates X by θ and uniformly scales it by $r > 0$. Then the $\oplus_\Delta$ and $\ominus_\Delta$ operations become identical to the polar dilation and erosion of Section 1.

Thus, both the translational and polar dilation/erosion result as special cases of the $\oplus_\Delta$ and $\ominus_\Delta$ operations by certain choices of the group G and the action Δ. The above very general approach, which we call "*group-action morphology*", achieves the desired unification, but without any further constraints it lacks sufficient structure to endow the resulting operations with useful properties. Hence, we next provide some constraints on Δ that are sufficient to enrich the $\oplus_\Delta$ and $\ominus_\Delta$ operations with invariance and distributivity properties. A sufficient such constraint will be for Δ_g to be an automorphism so that it is both a dilation (i.e., preserves unions) and an erosion (i.e., preserves intersections).

THEOREM 1 . *Let $B \subseteq G$ and define the operators δ_B, ε_B, α_B, $\phi_B \in \mathcal{O}$ by*

$$\delta_B(X) = X \oplus_\Delta B \quad , \quad \varepsilon_B(X) = X \ominus_\Delta B, \tag{11}$$

$$\alpha_B = \delta_B \varepsilon_B \quad , \quad \phi_B = \varepsilon_B \delta_B. \tag{12}$$

If the group action operator Δ_b is an automorphism for all $b \in B$, then:
(i) δ_B is a dilation (ii) ε_B is an erosion. (iii) $(\varepsilon_B, \delta_B)$ is an adjunction.
(iv) α_B is an opening and ϕ_B is a closing.

Proof: (i) If Δ_b is an automorphism, then it is a dilation and an erosion and hence distributes over $\cup$ and $\cap$. Then for any set collection $\{X_j : j \in J\}$

$$\delta_B(\bigcup_{j \in J} X_j) = \bigcup_{b \in B} \Delta_b(\bigcup_{j \in J} X_j) = \bigcup_{b \in B} \bigcup_{j \in J} \Delta_b(X_j) = \bigcup_{j \in J} \bigcup_{b \in B} \Delta_b(X_j) = \bigcup_{j \in J} \delta_B(X_j).$$

(ii) Similarly for ε_B by replacing $\cup$ with $\cap$. (iii) results from Propositions 2(iv) and 3(iii).
(iv) α_B and ϕ_B are increasing because they are the composition of two increasing operators. Since $(\varepsilon_B, \delta_B)$ is an adjunction, by Proposition 2 α_B is anti-extensive and $\alpha_B \alpha_B = (\delta_B \varepsilon_B \delta_B) \varepsilon_B = \delta_B \varepsilon_B = \alpha_B$. Thus α_B is idempotent. Hence α_B is an opening. Similarly, ϕ_B is a closing. □

In translational and polar morphology we have the serial laws: $(X \oplus A) \oplus B = X \oplus (A \oplus B)$ and $(X \ominus A) \ominus B = X \ominus (A \oplus B)$. Similar properties can be shown in group-action morphology:

PROPOSITION 4 . *Let $A, B \subseteq G$ and define $A \star B = \{a \star b : a \in A,\ b \in B\}$. If Δ_g is an automorphism for all $g \in A \cup B$, then*

$$(X \oplus_\Delta A) \oplus_\Delta B = X \oplus_\Delta (B \star A) \quad \text{and} \quad (X \ominus_\Delta A) \ominus_\Delta B = X \ominus_\Delta (A \star B). \tag{13}$$

In group-action morphology, the equivalent of "translation-invariance" for an image transformation Ψ is formulated as follows: Consider some $H \subseteq G$ and define its corresponding $\mathcal{H} \subseteq \mathcal{G}$ by $\mathcal{H} = \{\Delta_g : g \in H\}$. Then Ψ is called *$\mathcal{H}$-invariant* provided that

$$\Psi(\Delta_g(X)) = \Delta_g[\Psi(X)] \ , \ \forall g \in H \ , \ \forall X \subseteq S.$$

The set of all $g \in G$ for which $\Psi \Delta_g = \Delta_g \Psi$ is a subgroup of G. Therefore, when we talk about $\mathcal{H}$-invariance of an operator Ψ, we will always assume that H is a subgroup of G and thus $\mathcal{H}$ is a subgroup of $\mathcal{G}$. If the mapping $\xi = \Xi|_H : g \mapsto \Delta_g$ is defined on a subgroup H, then it is a group homomorphism between H and $\mathcal{H}$; further, if $Ker(\xi) = \{e\}$, then ξ is a bijection and $\mathcal{H}$ becomes isomorphic to H. We shall call Ψ

$\mathcal{G}$-invariant if $\mathcal{H} = \mathcal{G}$. For any subgroup $\mathcal{H} \subseteq \mathcal{G}$, the set of $\mathcal{H}$-invariant operators is closed under composition and contains *id*.

To see when the $\oplus_\Delta$ and $\ominus_\Delta$ operations are $\mathcal{H}$-invariant we need the following definitions: The *centralizer* of an element $g \in G$ is the subgroup $C(g) = \{a \in G : g \star a = a \star g\}$. The centralizer of a subgroup H of G is the subgroup

$$C(H) = \bigcap_{g \in H} C(g) = \{a \in G : g \star a = a \star g \; \forall q \in H\}$$

If $H = G$, $C(G)$ is a normal subgroup of G and is called the *center* of G. Obviously, if G is commutative, then $C(G) = G$.

THEOREM 2 . *Let H be a subroup of G and $\mathcal{H} = \{\Delta_g : g \in H\}$ a subgroup of $\mathcal{G} = \{\Delta_g : g \in G\}$. If $B \subseteq H \cap C(H)$ and $\mathcal{H}$ is a group of automorphisms, then the group-action dilation δ_B, the erosion ε_B, the opening α_B, and the closing ϕ_B are all $\mathcal{H}$-invariant operators.*

Proof: Since $B \subseteq H$, it follows from Theorem 1 that $\delta_B, \varepsilon_B, \alpha_B$, and ϕ_B are, respectively, a dilation, erosion, opening and closing. Since also $B \subseteq C(H)$ and Δ_g, $g \in H$, is an automorphism, δ_B is $\mathcal{H}$-invariant because for all $g \in H$

$$\begin{aligned} \delta_B(\Delta_g(X)) &= \bigcup_{b \in B} \Delta_b[\Delta_g(X)] = \bigcup_{b \in B} \Delta_{b \star g}(X) = \bigcup_{b \in B} \Delta_{g \star b}(X) \\ &= \bigcup_{b \in B} \Delta_g[\Delta_b(X)] = \Delta_g[\bigcup_{b \in B} \Delta_b(X)] = \Delta_g(\delta_B(X)). \end{aligned}$$

Similarly for ε_B by replacing $\cup$ with $\cap$ and Δ_b with $\Delta_{b^{-1}}$. The $\mathcal{H}$-invariance of α_B and ϕ_B follows from the closure of $\mathcal{H}$-invariant operators under composition. □

Next we provide a sufficient condition under which the group action becomes an automorphism, a property quite important for all the previous results. This condition requires that Δ be constructed from *point mappings*. We show this first for an arbitary operator Ψ:

PROPOSITION 5 . *Let $\Psi : \mathcal{P}(S) \to \mathcal{P}(S)$. Then Ψ is an automorphism on $\mathcal{P}(S)$ if and only if there exists a bijective point mapping $\sigma : S \to S$ such that*

$$\Psi(X) = \{\sigma(x) : x \in X\} \;\; , \;\; X \subseteq S. \tag{14}$$

Proof: Let $\Psi \in \mathcal{O}$ be constructed from (14) with σ being a bijection. Since $\sigma(x) \in \Psi(X) \iff x \in X$, it follows that $\Psi(X) = \Psi(Y) \iff X = Y$. Hence Ψ is a bijection. Further, from (14) and the definition of $\cup$ it follows that Ψ distributes over union and hence is a dilation. Since Ψ is a bijection and a dilation, by Proposition 1 it is an automorphism. Conversely, let Ψ be an automorphism. Since $\mathcal{P}(S)$ is an atomic lattice with a lowest element ($\emptyset$), each automorphism maps an atom of $\mathcal{P}(S)$ (i.e., each singleton $\{x\}$, $x \in S$) to another atom (see [6, p.21]). Thus, if $x \in S$, $\Psi(\{x\}) = \{y\}$ for some $y \in S$. This action of Ψ on the singleton sets induces a function $\sigma : S \to S$ defined by $\sigma(x) = y \iff \Psi(\{x\}) = \{y\}$. Since Ψ is a bijection, it follows that σ_g is a bijection on S. □

In a group action the only inherent property of the operators Δ_g that is related to the structure of automorphisms is the fact that Δ_g are bijections. By Proposition 5, Δ_g is an automorphism, if and only if it can be constructed from a point mapping, i.e., if there exists $\sigma_g : S \to S$ such that $\Delta_g(X) = \{\sigma_g(x) : x \in X\}$, $X \subseteq S$.

In Heijmans' [6] generalization of morphology *commutative* group structures G of $\mathbf{R}^2$ are first considered; such examples are the translation and the polar group [14]. These give rise to *commutative* groups of lattice automorphisms (i.e., bijective set operators that distribute over $\cup$ and $\cap$). Then dilation-type and erosion-type operations are defined as union and intersection, respectively, of a collection of automorphisms. Also in Heijmans & Ronse [7], although arbitrary automorphisms are briefly considered, the majority of results requires a group of commutative automorphisms. Group-action morphology is different from Heijmans', Roerdink's and Ronse's generalizations of translational morphology from two viewpoints: 1) In its most general form, group-action morphological operations are formed from union, intersection, or composition of less restricted and hence more general set operators; i.e., the group-action operators Δ_g are bijections but do not have to be automorphisms. 2) To endow the group-action morphology with some useful properties, we had to constrain Δ_g to be automorphisms, but even then they do not have to be commutative. For example, in the next section we provide a useful special case of group-action morphology where the underlying group of automorphisms is not commutative.

3 Affine Morphology

As a useful special case we introduce the *affine morphology*, which is defined from the general framework of group-action morphology by using $S = \mathbf{R}^2$ and as group G the set

$$G = \{(M,t) : M \in \mathbf{R}^{2\times 2}, det(M) \neq 0, t \in \mathbf{R}^2\} \tag{15}$$

equipped with the binary operation $(L,v)\star(M,t) = (LM, Lt+v)$. It can be easily shown that $(G,\star)$ is a *non-commutative* group. The group identity is $(I,0)$, where I is the identity matrix. The inverse of the element (M,t) is $(M,t)^{-1} = (M^{-1}, -M^{-1}t)$ where M^{-1} denotes matrix inverse. Toward defining an action of the group $(G,\star)$ on planar images $X \subseteq \mathbf{R}^2$, consider the affine set mappings

$$\Delta_{(M,t)}(X) = \{Mx + t : x \in X\} \ , \ X \subseteq \mathbf{R}^2. \tag{16}$$

Let $(M,t),(L,v) \in G$. Then

$$\Delta_{(L,v)}[\Delta_{(M,t)}(X)] = \{LMx + Lt + v : x \in X\} = \Delta_{(LM,Lt+v)}(X) = \Delta_{(L,v)\star(M,t)}(X)$$

Also $\Delta_{(I,0)}(X) = X$. Hence, Δ is a group action. Then $\mathcal{G} = \{\Delta_{(M,t)} : (M,t) \in G\}$ becomes the *affine group* $\mathcal{A}(2)$ of all invertible affine transformations of planar sets under composition. The 2×2 real matrix M performs rotation and scaling of the image X, while the vector t performs translation. Since $(I,0)$ is the only element in G that makes $\Delta_{(M,t)}$ the identity map, the action is faithful and the affine group $\mathcal{A}(2)$ is isomorphic to the group G.

For the affine, translational and polar morphology, note that all three correspond to group action operators $\Delta_{(M,t)}$ constructible from bijective point mappings

$$\sigma_{(M,t)}(x) = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = Mx + t \ , \ x \in \mathbf{R}^2, \tag{17}$$

where $x = (x_1, x_2)$ and $t = (t_1, t_2)$ are viewed both as points in the plane $\mathbf{R}^2$ with Cartesian coordinates and also as 2-dim real column vectors. The matrix M can always be written as

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} r_1\cos\theta_1 & -r_2\sin\theta_2 \\ r_1\sin\theta_1 & r_2\cos\theta_2 \end{bmatrix} \quad ; \quad \begin{matrix} a,b,c,d \in \mathbf{R} \\ r_1, r_2 \geq 0, \ \theta_1,\theta_2 \in [0, 2\pi) \end{matrix} \tag{18}$$

where (r_1, θ_1) are the polar coordinates of the point (a,c) and $(r_2, \theta_2 + \pi/2)$ are the polar cordinates of the point (b,d).

By constraining M or t in the general affine image transformation, several known classes of image transformations and can be derived, whose corresponding groups $\mathcal{H}$ are subgroups of $\mathcal{G} = \mathcal{A}(2)$. (We use the notation of Section 2.) If $M = I$, then $\Delta_{(M,t)}(X)$ becomes a *translation* $\tau_t(X) = X + t$ of the set X along the vector t. The set $\mathcal{H} = \{\tau_t : t \in \mathbf{R}^2\}$ under composition $\tau_t\tau_v = \tau_{t+v}$ becomes a commutative subgroup of $\mathcal{A}(2)$, called the *translation group* $\mathcal{T}(2)$. Translational morphology is created by the action of $\mathcal{T}(2)$ on $\mathcal{P}(\mathbf{R}^2)$.

In general, $det(M) = r_1 r_2 \cos(\theta_1 - \theta_2)$. If M is *orthonormal*, i.e., if $MM^T = I$ where $(\cdot)^T$ denotes transpose, then it follows that $r_1 = r_2 = 1$ and $|det(M)| = 1$. Let $i = 0$ and let M be orthonormal. Then, if $det(M) = 1$ ($\Longleftrightarrow \theta_1 = \theta_2 = \theta$), $\Delta_{(M,t)}$ becomes a counterclockwise *rotation* about the origin by an angle θ, denoted as

$$\lambda_\theta(X) = \{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : (x_1, x_2) \in X\} \tag{19}$$

whereas, if $det(M) = -1$ ($\Longleftrightarrow \theta_1 = \theta$ and $\theta_2 = \theta + \pi$), then $\Delta_{(M,t)}$ becomes a *reflection* about an axis. Thus both the rotation and the reflection can be represented by orthonormal (linear) transformations and are elements of the *orthogonal group* in two dimensions. The set of rotations is the *special orthogonal group* $\mathcal{SO}(2)$. If $r_1 = r_2 = 1$, then Δ is an *isometry*, i.e., a composition of an orthonormal transformation and a translation. The set of isometries is a group under composition, the *Euclidean group* $\mathcal{E}(2)$ of planar rigid motions.

If $t = 0$, $r_1 = r_2 = r$, and $\theta_1 = \theta_2 = 0$, then Δ becomes a *homothety* $\mu_r(X) = \{rx : x \in X\}$. The homothety performs a uniform contraction of X if $r < 1$ or a uniform expansion if $r > 1$. The set $\{\mu_r : r > 0\}$ under composition $\mu_r\mu_s = \mu_{rs}$ becomes a commutative subgroup of $\mathcal{A}(2)$, called the *uniform scaling group* $\mathcal{US}(2)$. Let $\nu_{(r,\theta)} = \mu_r\lambda_\theta = \lambda_\theta\mu_r$ be a composition of a homothety and rotation. The set $\mathcal{H} = \{\nu_{(r,\theta)} : r > 0, \theta \in [0, 2\pi)\}$ under composition $\nu_{(r,\theta)}\nu_{(\rho,\phi)} = \nu_{(r\rho,\theta+\phi)}$ becomes a commutative subgroup of $\mathcal{A}(2)$, isomorphic to the external direct product group $\mathcal{SO}(2) \times \mathcal{US}(2)$, and also isomorphic to the polar group of section 1. Polar morphology is created by the action of this $\mathcal{H}$ on $\mathcal{P}(\mathbf{R}^2)$.

By using as group of image transformations the general affine group $\mathcal{A}(2)$ or any of its subgroups above, we define the *affine dilation* of X by $B \subseteq G$ as

$$X \oplus_A B = \bigcup_{(M,t)\in B} \Delta_{(M,t)}(X) = \bigcup_{(M,t)\in B} \{Mx + t : x \in X\} = \delta_B(X). \tag{20}$$

The *affine erosion* of X by B is defined as

$$X \ominus_A B = \bigcap_{(M,t)\in B} \Delta_{(M,t)^{-1}}(X) = \bigcap_{(M,t)\in B} \{M^{-1}x - M^{-1}t : x \in X\} = \varepsilon_B(X). \tag{21}$$

Similarly, we define the *affine opening* $\circ_A$ and *affine closing* $\bullet_A$ of X by $B \subseteq G$ as

$$X \circ_A B = (X \ominus_A B) \oplus_A B = \alpha_B(X) \quad , \quad X \bullet_A B = (X \oplus_A B) \ominus_A B = \phi_B(X). \tag{22}$$

Note that the "structuring element" B is a collection of parameter pairs (M, t).

Obviously, translational and polar morphology are special cases of affine morphology. Three major differences between affine vs. translational and polar morphology are: 1) The acting group in affine morphology is *not* commutative, whereas the corresponding groups for translational and polar morphology are commutative. 2) The group action operator for affine morphology is not built from binary operations between planar points, as is the case for both the translational and polar morphology. 3) In translational and polar morphology the structuring element B is a subset of $\mathbf{R}^2$ and hence is a shape by itself; in contrast, the structuring element in affine morphology is a collection of parameters.

A common property shared by affine, translational and polar morphology is that in all three cases, due to Proposition 5, the operators $\Delta_{(M,t)}$ are automorphisms. Hence, based on Theorem 1, the affine dilation and erosion are increasing operators that distribute over $\cup$ and $\cap$, respectively. The affine opening (resp. closing) is increasing, anti-extensive (resp. extensive) and idempotent.

One potential application of affine morphology is the modeling of shape deformations that result in visual motion analysis when the 3-dim motion of a rigid object is projected on a 2-dim image plane. For example, in 3-dim coordinates (x, y, z), let the camera axis be along the z-axis and the image plane parallel to the (x, y)-plane. Then, if the 3-dim rigid object has X as its 2-dim projection (i.e., silhouette) on the image plane and its motion is a composition of a translation and a rotation both parallel to the image plane, then X undergoes a 2-dim isometry. A 3-dim translation along the z-axis corresponds to a uniform scaling (homothety) of X. Arbitrary 3-dim translations combined with a rotation around the z-axis incur a composition of isometry and scaling of X. Thus, many types of 3-dim rigid motion incur an affine transformation of the projected 2-dim silhouette. A parallel superposition of transformed silhouettes is also possible either when there are multiple objects projected on the plane or when there are multiple projections of the same object from different light sources. We can model such superpositions by using the affine dilation and erosion, or their combinations.

Another potential application is modeling of fractals. Specifically, let $\{w_n : n = 1, 2, ..., N\}$ be a finite set of *contractive* affine transformations; i.e., for some distance metric d in $\mathbf{R}^2$ there exists $0 \leq s < 1$ such that $d(w_n(x), w_n(y)) \leq s \cdot d(x, y)$ for all n. Then, on the space $\mathcal{K}(\mathbf{R}^2)$ of non-empty compact subsets of $\mathbf{R}^2$, the image transformation

$$W(X) = \bigcup_{n=1}^{N} w_n(X) \tag{23}$$

is a contraction with respect to the Hausdorff metric [10, 3]. Hence, by the contraction mapping theorem, there exists a unique fixed point F of W in $\mathcal{K}(\mathbf{R}^2)$, which can be obtained from iterations. That is, $F = \lim_{k\to\infty} W^{\circ k}(S)$, where $W^{\circ k}(S)$ is the k-fold self-composition of W and S is an arbitrary starting set in $\mathcal{K}(\mathbf{R}^2)$. $\{\mathbf{R}^2; w_n; n = 1, 2, ..., N\}$ is called in [3] an *iterated function system*, and F is called its *attractor*. In many cases F is a fractal image and can model shapes or textures occuring in natural scenes [3]. In the context of affine morphology, the mapping W is an affine dilation and F is the limit of this dilation when iterated. Thus, while in translational and polar morphology iterated dilations usually yield uninteresting results, iteration of affine dilations yields attractors that may have a remarkable degree of structure.

4 Affine Models for Signal Processing

In this section we extend the affine transformations to gray-level images and other signals, and we address the problem of modeling a signal by a superposition of affine signal transformations. Let $f : \mathbf{R}^d \to \mathbf{R}$, $d = 1, 2$, be a real-valued d-dim signal and consider the affine signal transformation

$$\Phi(f)(x) = rf(Mx + t) + c \tag{24}$$

where $M \in \mathbf{R}^{d\times d}$, $t \in \mathbf{R}^d$, and $r, c \in \mathbf{R}$. (Note: If $d = 2$, M is a 2×2 matrix and t is a vector. If $d = 1$, M and t are real numbers.) This corresponds to a 1-dim affine transformation of the amplitude range of the signal $y = f(x)$, i.e., $y \mapsto ry + c$, and to a d-dim affine transformation of its domain, i.e., $x \mapsto Mx + t$. If $G(f) = \{(x, f(x)) : x \in \mathbf{R}^d\}$ is the graph of f, then (24) is related to a $(d+1)$-dim affine set transformation of $G(f)$ that maps vertical lines into vertical lines. We create composite affine transformations

$$\Psi(f)(x) = V(\{\Phi_n(x) = r_n f(M_n x + t_n) + c_n : n = 1, 2, .., N\}) \tag{25}$$

of the signal f by superimposing several affine transformations Φ_n of f, defined by the indexed parameter tuples (M_n, t_n, r_n, c_n). The superposition is done via the value mapping $V : \mathcal{P}(\mathbf{R}) \to \mathbf{R}$. If $V()$ is sup() or inf(), then we obtain, respectively, an affine "dilation" or "erosion" of f. These terms become familiar from the following two examples. Let $M_n = I$ if $d = 2$ or $M_n = 1$ if $d = 1$, and let $r_n = 1$. Then, if $V() = \sup()$, $\Psi(f)$ becomes the gray-level morphological dilation [17] of f by the structuring function h

where $h(-t_n) = c_n$ and $h(x) = -\infty$ elsewhere. If $V() = \inf()$, then $\Psi(f)$ is the morphological erosion of f by h where $h(t_n) = -c_n$ and $-\infty$ elsewhere. However, in modeling signals with such a general superposition of affine transformations, if our goal is to estimate the model parameters, we need to sacrifice the generality and use a mathematically tractable superposition. Thus we focus only on the case where $V() = \sum_{n=1}^{N}$ and N is finite. Then we have

$$\Psi(f)(x) = c + \sum_{n=1}^{N} r_n f(M_n x + t_n) \tag{26}$$

where $c = \sum_n c_n$. We call this Ψ the *sum-affine* signal operator. Ψ is defined via the tuple $P = (M_1, ..., M_N, t_1, .., t_N, r_1, .., r_N, c)$ that contains $N(d^2 + d + 1) + 1$ real parameters. If $c = 0$ and $M_n = I$, then $\Psi(f)$ becomes the classical linear shift-invariant convolution of f with (the weighted impulse sum) $h(x) = \sum_{n=1}^{N} r_n \delta(x + t_n)$. If $c = 0$ and $M_n \neq I$, then Ψ is a linear but shift-varying signal operator. If $c \neq 0$, Ψ becomes nonlinear.

Now we formulate the problem of *approximately* modeling a signal f over a compact region K of $\mathbf{R}^d$ as the output of a sum-affine operator Ψ when the input is a signal h. We assume that f and h are known signals and follow a least squared error approach; i.e., we find the parameter tuple P that minimizes the error

$$E(P) = \int_K e^2(x)dx = \int_K [\Psi(h)(x) - f(x)]^2 dx = E(M_n, t_n, r_n, c). \tag{27}$$

If h is an input signal, then this is a system identification problem where an output signal f is modeled as a sum of affine amplitude-transformed and domain-transformed versions of the input h. Alternatively, this may be a signal modeling problem, where $h \neq f$ and h is some basis function, or some wavelet, or the impulse response of some system. If $f = h$, then our system Ψ becomes a *self-affine* model of f and has applications in modeling fractals as explained later. In general, minimizing E is a nonlinear minimization problem. However, it is linear with respect to the amplitude-scaling parameters r_n,c. Thus,

$$\frac{1}{2} \cdot \frac{\partial E}{\partial r_i} = \sum_{n=1}^{N} r_n \psi(n,i) + c\phi(i) - \psi(0,i), \quad i = 1, 2, .., N \tag{28}$$

$$\frac{1}{2} \cdot \frac{\partial E}{\partial c} = \sum_{n=1}^{N} r_n \phi(n) + c \int_K 1 dx - \int_K f(x)dx \tag{29}$$

where, if we use the notation $g_n(x) = h(M_n x + t_n)$,

$$\phi(n) = \int_K g_n(x)dx \ , \ \psi(n,i) = \int_K g_n(x) g_i(x)dx \ , \ \psi(0,i) = \int_K f(x) g_i(x)dx. \tag{30}$$

Setting $\partial E/\partial r_i = 0$ and $\partial E/\partial c = 0$ yields the set of $N + 1$ linear equations

$$\begin{aligned} \textstyle\sum_{n=1}^{N} r_n \psi(n,i) + c\phi(i) &= \psi(0,i), \quad i = 1, 2, .., N \\ \textstyle\sum_{n=1}^{N} r_n \phi(n) + c \int_K 1 dx &= \textstyle\int_K f(x)dx \end{aligned} \tag{31}$$

in the $N + 1$ unknowns r_n, c. By fixing M_n, t_n, we can solve these linear equations and obtain the optimal $r_n^\star$, $c^\star$, which are functions of M_n, t_n. Replacing then these optimal values back in E yields the new error

$$\begin{aligned} E^\star(M_n, t_n) &= E(M_n, t_n, r_n^\star, c^\star) \\ &= \int_K f^2(x)dx - \sum_{n=1}^{N} r_n^\star \int_K f(x) h(M_n x + t_n)dx - c^\star \int_K f(x)dx. \end{aligned} \tag{32}$$

Algorithm A: One approach to find the optimal $M_n^\star, t_n^\star$ is to assume that they have a finite and quantized range and perform an exhaustive search in this finite parameter space (M_n, t_n) of $N(d^2 + d)$ dimensions.

For example, if the domain of h contains the origin, then it is natural to constrain the translations t_n to be in the region K. The algorithm proceeds as follows: 1) For each allowed value (M_n, t_n) the signal correlations $\psi(n,i)$ and averages $\phi(n)$ are computed. 2) The set of linear equations (31) is solved to find the corresponding $r_n^\star, c^\star$. (If N is relatively small, then the closed-formula expressions that give $r_n^\star, c^\star$ as functions of M_n, t_n are simple and can be used directly.) 3) $E^\star$ is computed. Steps 1,2,3 are repeated throughout the entire parameter space and thus we find the optimal $M_n^\star, t_n^\star$ that minimize $E^\star$.

If the signals f, h have *discrete argument* $x \in \mathbf{Z}^d$, then the integration $\int_K$ becomes a summation over K in all the above formulas. Further, it is natural then to constrain all translations t_n to have integer values and be in K. However, the real parameters M_n (which contain information about domain-rotation and -scaling if $d = 2$ or about domain-scaling if $d = 1$) will be truncated during the computations since $Mx + t$ must always be integer.

A very special case of the sum-affine model Ψ and the previous model estimation approach has been used in [2] for multi-pulse linear predictive coding of speech ($d = 1$), where a short-time speech segment f is modeled by $\Psi(h)$ with $c = 0$, $M_n = 1$ (i.e., no domain scaling; only translation), and h being the impulse response of the linear prediction synthesis filter.

Algorithm B: An alternative approach to find the optimal parameter tuple P that minimizes E is to follow a traditional steepest descent approach. For simplicity, we describe this only for 1-dim signals ($d = 1$) where all M_n, t_n, r_n, c are real numbers and thus P is a real (column) vector. That is, choose a starting value P_1 for the parameter vector, and let P_k be its update at the k-th iteration step where

$$P_{k+1} = P_k - \lambda_k \nabla E(P_k), \quad k = 1, 2, 3, ... \tag{33}$$

and λ_k is a positive step size controlling the speed of convergence. The gradient of E is the (column) vector $\nabla E(P) = (.., \frac{\partial E}{\partial M_n}, .., \frac{\partial E}{\partial t_n}, .., \frac{\partial E}{\partial r_n}, .., \frac{\partial E}{\partial c})$, where

$$\frac{1}{2} \cdot \frac{\partial E}{\partial M_i} = r_i \cdot \int_K x h'(M_i x + t_i) \left[\sum_{n=1}^{N} r_n h(M_n x + t_n) + c - f(x) \right] dx \tag{34}$$

$$\frac{1}{2} \cdot \frac{\partial E}{\partial t_i} = r_i \cdot \int_K h'(M_i x + t_i) \left[\sum_{n=1}^{N} r_n h(M_n x + t_n) + c - f(x) \right] dx \tag{35}$$

where $h'(x) = dh(x)/dx$. If, for each k, λ_k is selected as the optimal solution to minimizing $E(P_k - \lambda \nabla E(P_k))$ subject to $\lambda \geq 0$, then $\lim_{k \to \infty} P_k = P_c$ where P_c is a stationary point, i.e., $\nabla E(P_c) = 0$. Let us make the second-order approximation

$$E(P) \approx E(P_k) + [\nabla E(P_k)]^T (P - P_k) + \frac{1}{2}(P - P_k)^T H (P - P_k) \tag{36}$$

where H is the Hessian matrix $[\partial^2 E / \partial p_i \partial p_j]$ evaluated at $P = P_k$, and p_i, $i = 1, 2, .., 3N + 1$ are components of P. Then, by using P_{k+1} from (33) in (36), it follows that $E(P_{k+1})$ is mimimized by the choice $\lambda_k = ||\nabla E||^2 / \nabla E^T H \nabla E$. In the extensive literature on optimization techniques, many other choises of λ_k or improvements of the steepest descent method can be found. The solutions P_c toward which these iterative optimization techniques converge could be a local or global minimum of E, but it could also be just a stationary point. However, in practice the usefulness of these solutions can only be evaluated empirically by performing experiments to see whether the solutions provided by such gradient descent methods are meaningful for the specific application.

APPLICATIONS:

I. If $N = 1$, then Ψ is a single affine transformation, and the algorithms A and B above provide least squares solutions in finding the parameters of Ψ. This has obvious applications in computer vision for motion detection since the M, t parameters provide information about the rotation/scaling and translation of an image object; in addition, the parameters r, c provide information about amplitude changes in the image brightness. Recovering the M, t parameters has also applications in recognizing deformed objects.

II. If $N > 1$, $f = h$, and f is a *binary image*, then by setting $c = 0$ and $r_n = 1$, $\Psi(f)$ becomes a sum of affine transformations of the binary image signal f. If Ψ were a union instead of a sum, this would be equivalent to an affine dilation of f. In addition, if all (M_n, t_n) were contractive mappings, $\Psi(f)$ would be what Barnsley [3] calls a "*collage*" of f made out of several patches, where each patch is a shrunk, rotated and translated version of f. Barnsley's collage theorem states that, the iterated mapping W in (23) with contraction factor $0 \leq s < 1$ converges to an attractor whose Hausdorff distance from a binary image X is no greater than the distance between X and $W(X)$ divided by $1 - s$. There has been considerable interest in finding the parameters of such a self-affine model of f. The approach followed in [3] is based on the Hausdorff metric, whose minimization with respect to the affine parameters is hard to analytically track using differential calculus. In this section, by replacing the union with a sum and the Hausdorff with the least squares metric, we provided two tractable algorithms to find the parameters of these self-affine models. Since these algorithms do not require f to be binary, another advantage of our approach is the fact that it easily extends to gray-level images simply by replacing the binary f with a gray-level image and by allowing the amplitude-scaling parameters r_n, c to vary freely. As such it can be applied to modeling fractal image textures by viewing them as the attractor of an iterated sum-affine model Ψ.

III. We also see some potential in applying the above sum-affine model to speech signals. One application results when $f \neq h$ and h is the impulse response of some speech synthesis system related to speech production. Then the given speech signal f could be modeled as a sum of affine domain-transformed and amplitude-scaled versions of h. This model generalizes the multipulse linear prediction scheme [2] because the latter involves only translations of the domain of h instead of general affine transformations. Another speech application arises when $f = h$; then $\Psi(f)$ becomes a self-affine model of f. Such a model could reflect the existence of self-similar structures in a short-time speech segment f; i.e., the possible fact that f, over a short time interval (in the order of one or a few pitch periods), is approximately a sum of affine-transformed versions of itself. Thus, given f, we can first approximate it by $\Psi(f)$ where the parameters of the sum-affine model Ψ are found via the previous least squares approaches. If these parameters are constrained so that Ψ is a contraction mapping,[1] then they are by themselves sufficient to approximately represent (i.e., synthesize) f by iteratively generating the attractor of Ψ. The maximum approximation error between f and the attractor of Ψ is determined by the modeling error between f and $\Psi(f)$, as is well known from the extensive literature on contraction mappings.

We are currently in the process of experimentally investigating the usefulness of these sum-affine models for speech and image processing.

Acknowledgements: This research was supported in part by the Army Research Office under Grant DAAL03-86-K-0171 to the Center for Intelligent Control Systems, and in part by the National Science Foundation under Grant MIPS-86-58150 with matching funds from Bellcore, DEC, Sun, and Xerox.

References

[1] S.-I. Amari, "Feature Spaces which Admit and Detect Invariant Signal Transformations", in *Proc. 4th Int. Joint Conf. on Pattern Recognition*, Kyoto, Japan, Nov. 1978.

---

[1] It is simple to show that by constraining the parameters r_n and M_n of Ψ, it is possible to make Ψ a contraction mapping. Then, in the complete signal space L_1 (or some other L_p space), Ψ will possess a fixed point (an *attractor*), which can be found as the limit of iterations of Ψ on an arbitrary starting signal.

[2] B. S. Atal and J. R. Remde, "A New Model of LPC Excitation for Producing Natural-Sounding Speech at Low Bit Rates", in *Proc. IEEE ICASSP '82*, pp. 614-617, Paris, 1982.

[3] M. F. Barnsley, *Fractals Everywhere*, Acad. Press, 1989.

[4] R. W. Brockett, "Gramians, Generalized Inverses, and the Least Squares Approximation of Optical Flow", in *Proc. IEEE Conf. Robot. & Autom.*, Mar. 1987.

[5] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche, und Isoperimetrie*, Berlin: Springer Verlag, 1957.

[6] H.J.A.M. Heijmans, "Mathematical Morphology: an Algebraic Approach", *CWI Newsletter*, no.14, Mar. 1987, Amsterdam.

[7] H.J.A.M. Heijmans and C. Ronse, "The algebraic basis of mathematical morphology. Part I: Dilations and erosions," CWI Rep. AM-R8807, Amsterdam, June 1988.

[8] W. C. Hoffman, "The Lie Algebra of Visual Perception", *J. Math. Psychology*, 3: 65-98, 1966.

[9] B. K. P. Horn, "Closed-form Solution of absolute orientation using unit quaternions", *J. Opt. Soc. Am. A*, 4, pp.629-642, Apr. 1987.

[10] J. E. Hutchinson, "Fractals and Self-Similarity", *Indiana Univ. Math. J.*, 30: 713-747, 1981.

[11] M. Lee, "Recovering the Affine Transformation of Images by Using Moments and Invariant Axes", in *Image Understanding and Machine Vision*, 1989 Tech. Digest Ser., vol.14, Washington, D.C.: Opt. Soc. Am., 1989, pp.2-5.

[12] P. Maragos and R. W. Schafer, "Morphological Systems for Multidimensional Signal Processing", *Proc. IEEE*, vol.78, pp. 690-710, April 1990.

[13] G. Matheron, *Random Sets and Integral Geometry*, NY: J. Wiley, 1975.

[14] J.B.T.M. Roerdink and H.J.A.M. Heijmans, "Mathematical Morphology for Structures without Translation Symmetry", *Signal Process.*, vol. 15, pp. 271-277, Oct. 1988.

[15] J. Serra, *Image Analysis and Mathematical Morphology*, NY: Acad. Press, 1982.

[16] J. Serra, Ed., *Image Analysis and Mathematical Morphology: Vol. 2 Theoretical Advances*, NY: Acad. Press, 1988.

[17] S. R. Sternberg, "Grayscale Morphology," *Comput. Vision, Graph., Image Proc.* **35**, pp.333-355, 1986.

[18] S. Ullman, "An Approach to Object Recognition: Aligning Pictorial Descriptions", A.I. Memo 931, A.I. Lab, M.I.T., Dec. 1986.