A DIGITAL FUNDAMENTAL GROUP

I. Y. KONG*
Department of Computer Science, Ohio University, Athens, Ohio 45701–2979, USA

Abstract—A group analogous to the fundamental group is defined for binary digital pictures based on almost arbitrary lattices and adjacency relations. This digital fundamental group has an immediate application to image processing—它可以 be used to give a precise statement of what it means for a three-dimensional image thinning algorithm to preserve topology. Given any suitably well-behaved digital picture, it is possible to construct a polyhedral set whose fundamental groups are naturally isomorphic to the digital fundamental groups of the digital picture. This result confirms that the digital fundamental group has been appropriately defined.

1. INTRODUCTION

The fundamental group is well known to topologists as one of the most basic topological invariants. This paper defines an analogous group for a large class of binary digital pictures—a digital fundamental group.

One application of the digital fundamental group is in the theory of three-dimensional image thinning. The natural way to specify that a thinning algorithm “preserves tunnels” is to say that the algorithm preserves digital fundamental groups. More on this in section 2.3.

Khalimsky has developed an interesting theory of fundamental groups (and higher homotopy groups) for a special kind of digital picture (see section 3.1 and [1]). However, his approach appears not to generalize easily to the kinds of digital picture that are most often used in computer graphics and image processing.

2. FUNDAMENTAL GROUPS OF CONVENTIONAL DIGITAL PICTURES

2.1 Conventional digital pictures

We write \( \mathbb{Z}^k \) for the set of all points in \( k \)-dimensional Euclidean space that have integer coordinates.

Two points in \( \mathbb{Z}^2 \) are said to be 8-adjacent if they are distinct and each coordinate of one differs from the corresponding coordinate of the other by at most 1; two points are 4-adjacent if they are 8-adjacent and differ in at most one of their coordinates. Two points in \( \mathbb{Z}^3 \) are said to be 26-adjacent if they are distinct and each coordinate of one differs from the corresponding coordinate of the other by at most 1; two points are 18-adjacent if they are 26-adjacent and differ in at most two of their coordinates; two points are 6-adjacent if they are 26-adjacent and differ in at most one coordinate. For \( n = 4, 8, 6, 18 \) or 26 an \( n \)-neighbor of a lattice point \( p \) is a point that is \( n \)-adjacent to \( p \).

A conventional digital picture is a quadruple \((V, m, n, B)\), where \( V = \mathbb{Z}^2 \) or \( V = \mathbb{Z}^3 \), \( B \subseteq V \), and where \((m, n) = (4, 8) \) or \((8, 4)\) if \( V = \mathbb{Z}^2 \) and \((m, n) = (26, 6) \) or \((18, 6)\) or \((6, 26)\) or \((6, 18)\) if \( V = \mathbb{Z}^3 \).

The conventional digital picture \( P = (V, m, n, B) \) is called two-dimensional or three-dimensional according as \( V = \mathbb{Z}^2 \) or \( V = \mathbb{Z}^3 \). The elements of \( V \) are called the points of \( P \). The points in \( B \) are called the black points of \( P \); the points in \( V - B \) are called the white points of \( P \). We call \( B \) and \( V - B \) the black point set and the white point set respectively.

Two black points in a conventional digital picture \((V, m, n, B)\) are said to be adjacent if they are \( m \)-adjacent. Two white points or a white point and a black point are said to be adjacent if they are \( n \)-adjacent. A point \( p \) is said to be adjacent to a set of points \( S \) if \( p \) is adjacent to some point in \( S \). A set of points \( S \) is said to be adjacent to a set of points \( T \) if some point in \( S \) is adjacent to some point in \( T \).

A \((V, m, n)\) digital picture (as in “a \((\mathbb{Z}^2, 8, 4)\) digital picture” or “a \((\mathbb{Z}^3, 6, 18)\) digital picture”) is a digital picture of form \((V, m, n, B)\) for some \( B \subseteq V \).

Let \( P \) be a conventional digital picture. A non-empty set of points of \( P \) is connected if it is not expressible as a union of two disjoint non-empty sets that are not adjacent to each other. A black component (white component) of \( P \) is a non-empty connected set \( C \) of black (white) points which is not adjacent to any black (white) point that is not in \( C \).

The complement of a conventional digital picture \( P = (V, m, n, B) \), denoted by \( P \), is the conventional digital picture \((V, m, n, V - B)\) obtained from \( P \) by switching the black and the white point sets and their associated adjacency relations.

2.2 The classical fundamental group

In classical topology a continuous curve in a set \( S \) from a point \( p \) in \( S \) back to itself is called a loop in \( S \) with base point \( p \). In other words a loop in \( S \) with base point \( p \) is a continuous map \( \gamma : [0, 1] \rightarrow S \) such that \( \gamma(0) = \gamma(1) = p \). A special case of a loop in \( S \) with base point \( p \) is the trivial loop \( \gamma(x) = p \) which remains fixed at \( p \).

If \( \gamma_1 \) and \( \gamma_2 \) are loops in \( S \) with the same base point, then the product of \( \gamma_1 \) and \( \gamma_2 \), written \( \gamma_1 \cdot \gamma_2 \), is the loop with base point \( p \) obtained by concatenating the loops \( \gamma_1 \) and \( \gamma_2 \) (and traversing the resulting loop twice as fast):

\[
\gamma_1 \cdot \gamma_2(x) = \begin{cases} 
\gamma_1(2x) & \text{if } 0 \leq x \leq \frac{1}{2} \\
\gamma_2(2x - 1) & \text{if } \frac{1}{2} \leq x \leq 1.
\end{cases}
\]

* On leave during the 1988–89 academic year. Current address: Department of Mathematics, City College of New York, CUNY, New York, NY 10031.
We write $\gamma^1$ for the loop obtained by traversing the loop $\gamma$ backwards—i.e., $\gamma^{-1}(x) = \gamma(1-x)$. We will also use the following notation: $\gamma^1 = \gamma^1_1, \gamma^n = \gamma^{n-1} \cdots \gamma^1$ (for integers $n > 1$), and $\gamma^n = (\gamma^{-1})^n$ (for integers $n > 1$).

Two loops $\gamma_1$ and $\gamma_2$ in a set $S$ that have the same base point $p$ are said to be equivalent in $S$ if there is a fixed base point deformation in $S$ that transforms $\gamma_1$ to $\gamma_2$; that is, $\gamma_1$ is equivalent in $S$ to $\gamma_2$ if there is a continuous map $h : [0, 1] \times [0, 1] \to S$ such that for all $0 < t < 1$ and $0 < s < t$:

1. $h(s, 0) = \gamma_1(s)$
2. $h(s, 1) = \gamma_2(s)$
3. $h(0, t) = h(1, t) = p$.

Let $[\gamma]_1$ denote the equivalence class of loops that are equivalent in $S$ to the loop $\gamma$. Then the product of two equivalence classes $[\gamma_1]_1$ and $[\gamma_2]_1$, written $[\gamma_1]_1 \cdot [\gamma_2]_1$, is defined to be $[\gamma_1 \cdot \gamma_2]_1$. (This is plainly a well-defined associative binary operation on equivalence classes.)

The fundamental group of a set $S$ with base point $p$, denoted by $\pi_1(S, p)$, is the group formed by the equivalence classes of loops in $S$ with base point $p$ under the product operation.

Examples. The fundamental group of a disk (with any base point) is a trivial group with just one element, because every loop in the disk is equivalent to a trivial loop. (In general, the fundamental group of a polyhedral set $S$ in 3-space is trivial if and only if $S$ has no holes or tunnels—see [2].)

The fundamental group of a circle (with any base point) is an infinite cyclic group. For every loop on the circle is equivalent to a trivial loop or to $\lambda^n$ (for some $n$) where $\lambda$ is a loop that goes around the circle just once, and $\lambda^n$ is not equivalent to $\lambda^m$ when $m \neq n$.

For more on the fundamental group see [3, Chs. 2, 3, and 4, Chs. 3, 4] or other textbooks on algebraic and geometric topology.

2.3 3D thinning and the need for a digital fundamental group

Thinning is an important pre-processing operation in pattern recognition. The purpose of thinning is to convert a digital picture $\mathcal{P}$ to one whose black point set is a "topologically equivalent skeleton" of the black point set of $\mathcal{P}$. A two-dimensional thinning algorithm is considered to be topologically sound if it satisfies the following criterion, which is essentially due to Stefanelli and Rosenfeld [5]:

**Criterion 2.3.1.** If a thinning algorithm converts $\mathcal{P}_1 = (\mathbb{Z}^3, m, n, B)$ to $\mathcal{P}_2 = (\mathbb{Z}^3, m, n, B - D)$ then

1. each black component of $\mathcal{P}_1$ contains exactly one black component of $\mathcal{P}_2$, and
2. each white component of $\mathcal{P}_2$ contains exactly one white component of $\mathcal{P}_1$.

This criterion is appropriate only in two dimensions. It is not stringent enough for 3D digital pictures, as it would permit any 3D digital picture whose black point set was a simple closed black digital curve—i.e., a connected set of black points with the property that each point in the set is adjacent to just two other points in the set—to be shrunk to a digital picture with just one black point.

In addition to conditions 1 and 2, a three-dimensional version of Criterion 2.3.1 must stipulate that the "tunnels" in the input and output digital pictures agree in number and position. More precisely, the thinning algorithm must preserve the "digital fundamental groups" of the input digital picture and its complement, where the digital fundamental group is some appropriately defined analog, for digital pictures, of the classical fundamental group. With this in mind, a three-dimensional version of Criterion 2.3.1 might be stated as follows.

**Criterion 2.3.2.** If a thinning algorithm converts the digital picture $\mathcal{P}_1 = (\mathbb{Z}^3, m, n, B)$ to $\mathcal{P}_2 = (\mathbb{Z}^3, m, n, B - D)$ then

1. each black component of $\mathcal{P}_1$ contains exactly one black component of $\mathcal{P}_2$, and
2. each white component of $\mathcal{P}_2$ contains exactly one white component of $\mathcal{P}_1$, and
3. for each point $p$ in $B - D$ (i.e., for each black point of $\mathcal{P}_2$), the inclusion map $i : B - D \to B$ induces a group isomorphism $i_* : \pi(\mathcal{P}_2, p) \to \pi(\mathcal{P}_1, p)$, and
4. for each point $q$ in $\mathbb{Z}^3 - B$ (i.e., for each white point of $\mathcal{P}_1$), the inclusion map $j : \mathbb{Z}^3 - B \to \mathbb{Z}^3 - (B - D)$ induces a group isomorphism $j_* : \pi(\mathcal{P}_1, q) \to \pi(\mathcal{P}_2, q)$.

where $\pi(\mathcal{P}, x)$ denotes the digital fundamental group of the digital picture $\mathcal{P}$ with base point $x$.

The main purpose of this paper is to discuss possible ways of defining the digital fundamental group $\pi(\mathcal{P}, x)$ used in conditions 3 and 4.

2.4 The digital fundamental group

Recall that the definition of the classical fundamental group depends on the notions of loop, product of loops and loop equivalence. In this section we formulate discrete analogs of these concepts, which we then use to define the digital fundamental group.

It is easy to define discrete analogs of loops and products of loops. Let $p$ be any black point of a digital picture $\mathcal{P}$. Define a black digital loop of $\mathcal{P}$ with base point $p$ to be a sequence $\{p_0, p_1, \cdots, p_n\}$ of black points of $\mathcal{P}$ where $p_0 = p_n$ and where each point $p_i$ is equal or adjacent to $p_{i-1}$ (1 $\leq i \leq n$). The product of two black digital loops $c_1 = \langle x = p_0, p_1, \cdots, p_m - x \rangle$ and $c_2 = \langle x = q_0, q_1, \cdots, q_n - x \rangle$, written $c_1 \cdot c_2$, is the black digital loop $\langle x = p_0, p_1, \cdots, p_m = x = q_0, q_1, \cdots, q_n = x \rangle$ obtained by concatenating $c_1$ and $c_2$.

It remains to formulate a discrete analog of loop equivalence.

Say that a black digital loop $\langle x = p_0, p_1, \cdots, p_m = x \rangle$ is in reduced form if $p_i \neq p_{i+1}$ for all $0 < i < m$. Define the reduced form of a black digital loop $c = \langle x = p_0, p_1, \cdots, p_m = x \rangle$ to be the longest black digital loop $\langle x, p_i, \cdots, p_i = x \rangle$ where $i$ is the smallest value of $i$ such that $p_i \neq x$, and where each of the
other \( i_k \) is the smallest value of \( i \) greater than \( i_{k-1} \) such that \( p_i \neq p_{i_k} \). In the special case where all the \( p_i \) are equal to \( x \), define the reduced form of \( c \) to be \( \langle x \rangle \). In other words, the reduced form of any black digital loop is obtained by removing all but one point from every set of consecutive equal points.

Let \( \mathcal{P} \) be a conventional digital picture and let \( K \) be any unit lattice square or unit lattice cube. (In 3-space the unit lattice squares are the faces of the unit lattice cubes.) Say that two black digital loops \( c_1 = \langle x = p_1, p_2, \ldots, p_m = x \rangle \) and \( c_2 = \langle x = q_1, q_2, \ldots, q_n = x \rangle \) in \( \mathcal{P} \) differ only in \( K \) if

1. \( m = n \), and
2. \( p_i \in K \) whenever \( q_i \in K \), and
3. \( p_i = q_i \) whenever \( q_i \notin K \).

In other words, \( c_1 \) and \( c_2 \) are said to differ only in \( K \) if \( c_1 \) and \( c_2 \) are the same when all points in \( K \) are identified with each other.

Say that two black digital loops \( c_1 \) and \( c_2 \) in a two- or three-dimensional conventional digital picture \( \mathcal{P} \) that have the same base point are directly equivalent if

1. \( c_1 \) and \( c_2 \) have the same reduced form, or
2. \( c_1 \) and \( c_2 \) differ only in a unit lattice square \( K \), or
3. (when \( \mathcal{P} \) is three-dimensional) \( c_1 \) and \( c_2 \) differ only in a unit lattice cube \( K \), provided that if \( \mathcal{P} \) is a \((\mathbb{Z}^3, 6, 26)\) digital picture then the cube \( K \) does not contain two white points that are diametrically opposite one another.

Say that two black digital loops \( c \) and \( c' \) of \( \mathcal{P} \) with the same base point \( p \) are equivalent in \( \mathcal{P} \) if there is a sequence of black digital loops \( c_1, \ldots, c_k \) with base point \( p \) such that \( c_i \) is directly equivalent to \( c \), \( c' \) is directly equivalent to \( c_k \) and \( c_{i+1} \) is directly equivalent to \( c_i \) for \( 1 \leq i < k \). In other words, equivalence is the transitive closure of direct equivalence.

It may not be immediately clear why in case 3 of the definition of direct equivalence we stipulated that "if \( \mathcal{P} \) is a \((\mathbb{Z}^3, 6, 26)\) digital picture then the cube \( K \) does not contain two white points that are diametrically opposite one another." The purpose of this requirement is to avoid a potential paradox which we now describe.

Fig. 1 shows five 5-point by 5-point planes that make up a 5-point by 5-point by 5-point cube in a \((\mathbb{Z}^3, 6, 26)\) digital picture. (So the point \( c \) is at the center of the cube.) The 1's and the points \( a, b, c, d, e \) and \( f \) are black points; the 0's are white points. Intuitively, the 0's "form a tunnel" in the cube. But if in case 3 of the definition of direct equivalence we omitted the above-mentioned stipulation then the black digital loop \( \langle a, b, c, d, e, f, a \rangle \) would be directly equivalent to the loop \( \langle a, b, c, a, a, a, a \rangle \), which is directly equivalent to the trivial loop \( \langle a \rangle \). This would be a paradox—the loop \( \langle a, b, c, d, e, f, a \rangle \) and the tunnel formed by the 0's are intuitively "linked", and so the loop should not be deformable to a one-point loop.

The set of black digital loops that are equivalent in \( \mathcal{P} \) to a black digital loop \( c \) is denoted by \([c]_p \). The product of two equivalence classes \([c_1]_p \) and \([c_2]_p \), written \([c_1]_p \cdot [c_2]_p \), is defined to be \([c_1 \cdot c_2]_p \). As in the continuous case, this is a well-defined associative binary operation on equivalence classes.

The digital fundamental group of a conventional digital picture is now defined in an analogous way to the classical fundamental group:

**Definition 2.4.1.** The digital fundamental group of a digital picture \( \mathcal{P} \) with base point \( p \), denoted by \( \pi(\mathcal{P}, p) \), is the group formed by the equivalence classes of black digital loops in \( \mathcal{P} \) with base point \( p \) under the \product operation.
3. FUNDAMENTAL GROUPS OF GENERAL DIGITAL PICTURES

3.1 Other kinds of digital picture

The definition of $\pi(\mathcal{P}, x)$ given in the previous section only applies to conventional digital pictures. However, many other kinds of digital pictures have been considered in the literature.

The 2D isometric hexagonal lattice was used in [6–9]. Both hexagonal and triangular lattices were considered in [10]. Kovalevsky [11] proposed a tessellation of three-space by truncated octahedra, which corresponds to a body-centered cubic lattice. In [12] Itoh, Waki et al. considered digital pictures on a face-centered cubic lattice.

On non-orthogonal lattices such as these the analogs of the 4-, 8-, 6-, 18- and 26-adjacency relations are the adjacency relations based on the Voronoi neighborhoods of the grid points [13]. The Voronoi neighborhood of a grid point $p$ in a 2D grid (3D grid) is the set of all points in the plane (in 3-space) that are at least as close to $p$ as in any other grid point. It is always a closed convex polygon or polyhedron. The Voronoi adjacency relations are the adjacency relations in which two grid points are adjacent if their Voronoi neighborhoods (i) share a vertex, (ii) share an edge, and (iii) share a face. However, these adjacency relations may not all be distinct.

A Voronoi adjacency relation in which each grid point is adjacent to just $n$ other grid points is referred to as the $n$-adjacency relation on the grid; two adjacent points are then said to be $n$-adjacent to each other and each is called an $n$-neighbor of the other. Note that the 4-, 8-, 6-, 18- and 26-adjacency relations on the square and cubic lattices are special cases of this. We call a straight line segment joining two $n$-neighbors an $n$-adjacency.

It is readily confirmed that there are just two Voronoi adjacency relations—12- and 18-adjacency—on the face-centered cubic lattice

$$\{ (x, y, z) \in \mathbb{Z}^3 | x + y + z = 0 \pmod{2} \},$$

and that there is just one Voronoi adjacency relation—14-adjacency—on the body-centered cubic lattice

$$\{ (x, y, z) \in \mathbb{Z}^3 | x = y = z \pmod{2} \}.$$

Khalimsky has introduced an entirely different kind of digital picture (see [14–16]), which has also been considered by Kovalevsky [17]. Khalimsky defines a topology on the set $\mathbb{Z}^n$ for every positive integer $n$; two lattice points $p$ and $q$ are adjacent if the two-point set $\{ p, q \}$ is connected with respect to that topology.

Following Khalimsky, a lattice point $p$ is a pure point if all of its coordinates are even numbers, or if all of its coordinates are odd numbers. Call the other lattice points mixed points. Then two lattice points in a two-dimensional Khalimsky digital picture are adjacent if and only if they are 8 adjacent or 26 adjacent and at least one of them is a pure point, or (in the 3D case) if they are 6-adjacent mixed points. Khalimsky’s digital pictures are mathematically well-behaved—see [18], for example.

Digital pictures of Khalimsky’s type can also be defined for non-orthogonal grids. Indeed, Kovalevsky’s construction of Khalimsky’s digital pictures in [17] generalizes easily to Sklansky and Kibler’s arbitrary mosaics [19] and to the analogs of such mosaics in three and higher dimensions.

As we mentioned in the introduction, Khalimsky has developed a theory of fundamental groups (and higher homotopy groups) for his digital pictures [1]. However, that theory does not appear to generalize to other kinds of digital picture; in particular, it is not applicable to the conventional digital pictures.

We shall give a more general definition of the digital fundamental group. The new definition will be equivalent to Definition 2.4.1 for conventional digital pictures (up to a natural group isomorphism). But unlike Definition 2.4.1 it will also be applicable to the kinds of digital pictures we have just described.

3.2 General digital pictures: Terminology and notation

We now generalize the terminology and notation introduced in Section 2.1.

A digital picture space is a triple $(V, \beta, \omega)$, where $V$ is the set of grid points in a 2D or 3D grid and each of $\beta$ and $\omega$ is a set of straight line segments joining pairs of points in $V$. A digital picture space is called two-dimensional or three-dimensional according as the grid is two-dimensional or three-dimensional. A line segment in $\beta$ is called a $\beta$-adjacency. Similarly, a line segment in $\omega$ is called an $\omega$-adjacency. If $p$ and $q$ are the endpoints of a $\beta$-adjacency ($\omega$-adjacency) we say $p$ is $\beta$-adjacent ($\omega$-adjacent) to $q$. If a point $p$ is $\beta$-adjacent ($\omega$-adjacent) to a point $q$ then we say $p$ is a $\beta$-neighbor ($\omega$-neighbor) of $q$.

For brevity a digital picture space will sometimes be referred to simply as a space.

A digital picture is a quadruple $(V, \beta, \omega, B)$, where $(V, \beta, \omega)$ is a digital picture space and $B$ is a subset of $V$. The digital picture is called two-dimensional or three-dimensional according as the digital picture space is two- or three-dimensional. We say $(V, \beta, \omega, B)$ is a digital picture on the space $(V, \beta, \omega)$. The points in $B$ are called the black points of (or in) the digital picture. The points in $V - B$ are called white points of $\omega$ in the digital picture.

Two black points in the digital picture $\mathcal{P} = (V, \beta, \omega, B)$ are said to be adjacent (with respect to $\mathcal{P}$) if they are $\beta$-adjacent. Two white points or a white point and a black point are said to be adjacent if they are $\omega$-adjacent.

A $\beta$-adjacency that joins two black points is called a black adjacency. An $\omega$-adjacency that joins two white points is called a white adjacency.

The conventional digital picture $(\mathbb{Z}^2, m, n, B)$ is identified with the digital picture $(\mathbb{Z}^2, \beta, \omega, B)$ in which $\beta$ and $\omega$ are respectively the set of all $m$-adjacencies and the set of all $n$-adjacencies of $\mathbb{Z}^2$.

We now generalize the notation $(\mathbb{Z}^2, m, n, B)$ which
we have been using for conventional digital pictures. If $\beta$ is the set of all $m$-adjacencies of $V$ and $\omega$ is the set of all $n$-adjacencies of $V$ then the digital picture space $(V, \beta, \omega)$ may be denoted by $(V, m, n)$ and the digital picture $(V, \beta, \omega, \rho)$ may be denoted by $(V, m, n, \rho)$. For example, if $V$ is the set of grid points of a face-centered cubic grid then $(V, 12, 18)$ is a valid digital picture space.

The complement of a digital picture space $(V, \beta, \omega)$ is the space $(V, \omega, \beta)$. The complement of a digital picture $P = (V, \beta, \omega, B)$, written $\overline{P}$, is the digital picture $(V, \omega, \beta, V - B)$.

Connected sets of points and black and white components in a general digital picture $P$ are defined in the same way as they are defined in a conventional digital picture—see section 2.1.

An isomorphism of a digital picture space $\Sigma_1 = (V_1, \beta_1, \omega_1)$ to a digital picture space $\Sigma_2 = (V_2, \beta_2, \omega_2)$ is a homeomorphism $h$ of the Euclidean plane (2D case) or Euclidean 3-space (3D case) to itself such that $h$ maps $V_1$ onto $V_2$, each $\beta_1$-adjacency onto a $\beta_2$-adjacency and each $\omega_1$-adjacency onto a $\omega_2$-adjacency, and $h^{-1}$ maps each $\beta_2$-adjacency onto a $\beta_1$-adjacency and each $\omega_2$-adjacency onto an $\omega_1$-adjacency.

An isomorphism of a digital picture $P_1 = (V_1, \beta_1, \omega_1, B_1)$ to a digital picture $P_2 = (V_2, \beta_2, \omega_2, B_2)$ is an isomorphism of the digital picture space $(V_1, \beta_1, \omega_1)$ to the digital picture space $(V_2, \beta_2, \omega_2)$ that maps $B_1$ onto $B_2$.

Isomorphic spaces and digital pictures are essentially indistinguishable in the study of digital topology.

3.3 A more general digital fundamental group

We write $E^2$ for the Euclidean plane and $E^3$ for Euclidean 3 space.

Let $P$ be a black point of a digital picture $\mathcal{P}$. A loop of $\mathcal{P}$ with base point $p$ is a continuous map $\gamma : [0, 1] \to E^n$, where $n = 2$ or 3 according as $\mathcal{P}$ is two-dimensional or three-dimensional, such that $\gamma(0) = \gamma(1) = p$ and there exists a positive integer $k$ such that for all non-negative integers $i < k$.

1. $\gamma(i/k)$ is a black point of $\mathcal{P}$
2. $\gamma(i/k)$ is equal to or adjacent to $\gamma(i + 1/k)$
3. $\gamma$ is linear on the closed interval $[i/k, (i + 1)/k]$; that is, $\gamma((i + h)/k) = (1 - h)\gamma(i/k) + h\gamma((i + 1)/k)$ for $0 \leq h \leq 1$.

Note that unlike a black digital loop, a loop of $\mathcal{P}$ is in fact a loop in the sense of section 2.2—it is a loop in the union of the set of black adjacencies of $\mathcal{P}$. We shall write $\mathcal{P}_B$ for the union of the set of black adjacencies of $\mathcal{P}$. This a loop of $\mathcal{P}$ is a special kind of loop in $\mathcal{P}_B$.

We now adapt the notion of loop equivalence given in section 2.2 to give an appropriate notion of equivalence for loops of a digital picture. We avoid "pathological" spaces that are incompatible with our definition of equivalence by imposing some restrictions on the sets $\beta$ and $\omega$ of a space $(V, \beta, \omega)$.

Definition 3.3.1. A digital picture space $(V, \beta, \omega)$ is said to be regular if it satisfies the following conditions:

1. No $\beta$-adjacency or $\omega$-adjacency passes through any point in $V$ other than its endpoints.
2. No $\beta$-adjacency meets an $\omega$-adjacency with which it does not share an endpoint. (This means that no $\beta$-adjacency 'crosses' an $\omega$-adjacency.)

Slightly different definitions of loop equivalence are needed for two- and three-dimensional digital pictures. The definition for 2D digital pictures is as follows:

Definition 3.3.2. Let $P = (V, \beta, \omega, B)$ be a digital picture on a regular two-dimensional digital picture space. Two loops $\gamma_1$ and $\gamma_2$ in $\mathcal{P}_B$ are said to be equivalent in $\mathcal{P}$ if $\lambda_1$ and $\lambda_2$ are equivalent in $E^2 - (V - B)$.

In other words, two loops of a 2D digital picture $P$ on a regular digital picture space and more generally two loops in $\mathcal{P}_B$ are equivalent if one loop can be transformed to the other by a fixed base point continuous deformation that avoids all white points of $P$.

In 3-space there is enough room to maneuver a loop around any number of white points, so this definition fails. Two loops of a 3D digital picture $P$ on a regular space and more generally two loops in $\mathcal{P}_B$ will only be called equivalent if one loop can be transformed to the other by a fixed base point continuous deformation that avoids all white adjacencies of $\mathcal{P}$.

Definition 3.3.3. Let $P = (V, \beta, \omega, B)$ be a digital picture on a regular three-dimensional space, and let $A$ be the union of the set of white adjacencies of $P$. Two loops $\lambda_1$ and $\lambda_2$ in $\mathcal{P}_B$ are said to be equivalent in $\mathcal{P}$ if $\lambda_1$ and $\lambda_2$ are equivalent in $E^3 - A$.

Every loop in $\mathcal{P}_B$ with base point in $B$ is equivalent in $\mathcal{P}$ to a loop of $\mathcal{P}$. For $\lambda \in \mathcal{P}_B$ let $[\lambda]_B$ denote the set of all loops of $\mathcal{P}$ that are equivalent in $\mathcal{P}$ to the loop $\lambda$. If the loops $\lambda$ and $\chi$ in $\mathcal{P}_B$ have the same base point, then the product of $[\lambda]_B$ and $[\chi]_B$, written $[\lambda]_B \cdot [\chi]_B$, is defined to be $[\lambda \cdot \chi]_B$. This is a well defined associative binary operation.

The digital fundamental group of $\mathcal{P}$ with base point $p$ can now be redefined as follows:

Definition 3.3.4. Let $\mathcal{P}$ be a digital picture on a regular digital picture space. The digital fundamental group of $\mathcal{P}$ with base point $p$, denoted by $\pi(\mathcal{P}, p)$, is the group formed by all the sets $[\lambda]_B$ in which $\lambda$ is a loop in $\mathcal{P}_B$ based at $p$, under the product operation.

We claimed earlier that our more general definition of the digital fundamental group (i.e., Definition 3.3.4) would be equivalent to Definition 2.4.1 up to a natural group isomorphism whenever $\mathcal{P}$ happened to be a conventional digital picture. This can be deduced from [20, Proposition 7.9.1].

Digital fundamental groups are of course invariant under isomorphism of digital pictures. Indeed, if $f$ is any isomorphism of a digital picture $P_1$ to a digital picture $P_2$, then for each black point $p$ in $P_1$, $f$ induces a group isomorphism of $\pi(P_1, p)$ to $\pi(P_2, f(p))$.

4. STRONGLY NORMAL DIGITAL PICTURE SPACES

4.1 Continuous analogs

The goal of section 4 is to identify a large class of digital picture spaces—the strongly normal spaces—

1 In the sense of section 2.2.
on which the above definition of the digital fundamental group is intuitively correct.

Given an n-dimensional digital picture \( P = (V, \beta, \omega, B) \), there may be a polyhedral set \( C(P) \supseteq B \) (a "continuous analog" of \( P \)) that has all of the following properties:

1. All black points and all black adjacencies of \( P \) are contained in \( C(P) \).
2. All white points and all white adjacencies of \( P \) are contained in \( E^w - C(P) \).
3. Each connected component of \( C(P) \) meets \( Z^n \) in a black component of \( P \).
4. Each connected component of \( E^w - C(P) \) meets \( Z^n \) in a white component of \( P \).
5. The boundary of a connected component \( X \) of \( C(P) \) meets the boundary of a connected component \( Y \) of \( E^w - C(P) \) if and only if there is a black point in \( X \) that is adjacent to a white point in \( Y \).
6. For each black point \( p \) in \( P \), the inclusion of the black points and black adjacencies of \( P \) in \( C(P) \) induces an isomorphism of the digital fundamental group \( \pi_1(C(P), p) \) to the (classical) fundamental group \( \pi_1(C(P), p) \).
7. For each white point \( q \) in \( P \), the inclusion of the white points and white adjacencies of \( P \) in \( E^w - C(P) \) induces an isomorphism of the digital fundamental group \( \pi_1(E^w - C(P), q) \) to the fundamental group \( \pi_1(E^w - C(P), q) \).

Observe that if a digital picture \( P \) has a continuous analog satisfying the above conditions, then so does every digital picture that is isomorphic to \( P \).

When such a continuous analog exists, conditions 6 and 7 are evidence that the digital fundamental group has been appropriately defined. A strongly normal space has the property that every digital picture on the space has such a continuous analog. There are in fact many other spaces with the same property. However, the class of strongly normal spaces and the spaces isomorphic to them already include most of the digital picture spaces that have been considered in the literature.

4.2. Definition of a strongly normal space

A two-dimensional or three-dimensional digital picture space \( \mathcal{S} = (V, \beta, \omega) \) is strongly normal if it is regular and also satisfies all of the following conditions:

1. \( V = \mathbb{Z}^2 \) or \( \mathbb{Z}^3 \).
2. Every 4-adjacency (2D case) or every 6-adjacency (3D case) is both a \( \beta \)-adjacency and an \( \omega \)-adjacency.
3. All \( \beta \)-adjacencies and \( \omega \)-adjacencies are 8-adjacencies (2D case) or 26-adjacencies (3D case).
4. In any given unit lattice square either both diagonals are \( \beta \)-adjacencies or both diagonals are \( \omega \)-adjacencies or one of the diagonals is both a \( \beta \)-adjacency and an \( \omega \)-adjacency.
5. In the case \( V = \mathbb{Z}^2 \) every digital picture \( P \) on \( \mathcal{S} \) has the property that whenever a black component of \( P \) is either \( \beta \)-adjacent or \( \omega \)-adjacent to a white component of \( P \), the black component is also 6-adjacent to the white component.

Observe that the complement of a strongly normal space is strongly normal.

Condition 1 is not as restrictive as it might seem to be, because a digital picture space based on a non-orthogonal grid may still be isomorphic to a strongly normal space.

Regarding condition 4, note that if both diagonals of a unit lattice square are \( \beta \)-adjacencies then by regularity neither diagonal can be an \( \omega \)-adjacency, and vice versa. Equivalently, if one of the diagonals is both a \( \beta \)-adjacency and an \( \omega \)-adjacency then the other diagonal is neither a \( \beta \)-adjacency nor an \( \omega \)-adjacency. A corollary of this observation is that in all strongly normal spaces with \( \beta = \omega \) exactly one diagonal of each unit lattice square is an adjacency.

Though condition 5 is stated only for three-dimensional digital picture spaces, it is easy to verify (see [20]) that strongly normal two-dimensional spaces have an analogous property:

**Proposition 4.2.1.** If a two-dimensional digital picture space \( \mathcal{S} = (\mathbb{Z}^2, \beta, \omega) \) is strongly normal, then in any digital picture on \( \mathcal{S} \) a black component and a white component that are \( \beta \)-adjacent or \( \omega \)-adjacent are also 4-adjacent.

This result and condition 5 imply that a black component and a white component of a digital picture on a strongly normal space \( (V, \beta, \omega) \) are \( \beta \)-adjacent if and only if they are \( \omega \)-adjacent.

4.3. Examples of strongly normal spaces

The spaces \((\mathbb{Z}^2, 8, 4), (\mathbb{Z}^3, 4, 8), (\mathbb{Z}^3, 6, 26), (\mathbb{Z}^3, 26, 6), (\mathbb{Z}^3, 6, 18) \) and \((\mathbb{Z}^4, 18, 6)\) are all strongly normal.

The Khalimsky spaces \((\mathbb{Z}^n, \beta, \omega)\)—in which \( \beta = \omega \) is the \( n \)-dimensional Khalimsky adjacencies—are strongly normal for \( n = 2 \) or 3.

Each of the following spaces is isomorphic to a strongly normal space:

1. \( (V, 6, 6) \) where \( V \) - grid points of the 2D isotropic hexagonal grid
2. \( (V, 14, 14) \) where \( V \) - grid points of the 3D body-centered cubic grid
3. \( (V, 12, 12) \) where \( V \) - grid points of the 3D face-centered cubic grid
4. \( (V, 12, 18) \) where \( V \) - grid points of the 3D face-centered cubic grid
5. \( (V, 18, 12) \) where \( V \) - grid points of the 3D face-centered cubic grid

The first of these five spaces is isomorphic to the strongly normal space \((\mathbb{Z}^n, \alpha, \alpha)\), where \( \alpha \) contains all 4-adjacencies and all 8-adjacencies that join a lattice point \((x, y)\) to \((x + 1, y - 1)\). It is also quite easy to find strongly normal isomorphs of the other four spaces. In fact it is only necessary to do this for the second, third and fourth spaces, since the fifth space is just the complement of the fourth. To describe isomorphs of the second, third and fourth spaces, let \( \alpha_5 \) be the set of 18-adjacencies in the 3D cubic grid whose endpoints \((x, y, z)\) and \((x', y', z')\) satisfy \( x + y + z = x' + y' + z'\). Also, let \( \alpha_6 \) be the set of 26-adjacencies
that are not 6-adjacencies but whose endpoints \((x, y, z)\) and \((x', y', z')\) satisfy \(|x + y + z| - |x' + y' + z'| - 1\), and let \(a_0\) be the set of 26-adjacencies whose endpoints \((x, y, z)\) and \((x', y', z')\) satisfy \(|x + y + z| - |x' + y' + z'| \geq 2\). Finally, let \(a_0\) be the set of all 6-adjacencies in the 3D cubic grid. Then the three spaces \((\mathbb{Z}^2, a_0 \cup a_1, a_0 \cup a_1, a_0 \cup a_2)\) and \((\mathbb{Z}^2, a_0 \cup a_1, a_0 \cup a_1, a_0 \cup a_2)\) are strongly normal, and are respectively isomorphic to the second, third and fourth spaces on the above list.

Thus most of the digital picture spaces that have been used in the literature on digital geometry are either strongly normal or isomorphic to a strongly normal space.

4.4 A discrete definition of the digital fundamental group for strongly normal spaces

It turns out that a slight modification of the definition of equivalence between black digital loops will make Definition 2.4.1 correct for all strongly normal spaces.

Let \(\mathcal{P}\) be a digital picture on a strongly normal space and let \(K\) be any unit lattice square or unit lattice cube. Say that two black digital loops \(c_1 = \langle x = p_0, p_1, p_2, \ldots, p_m = x' \rangle\) and \(c_2 = \langle x = q_0, q_1, q_2, \ldots, q_n = x' \rangle\) in \(\mathcal{P}\) are \(K\)-equivalent if

1. \(m = n\), and
2. \(p_i \in K\) whenever \(q_i \in K\), and
3. \(p_i \in K\) whenever \(q_i \notin K\) or \(q_i \in K\) or \(q_{i+1} \notin K\)

Say that two black digital loops with the same base point are \textit{immediately equivalent} if they have the same reduced form, or if they are \(K\)-equivalent for some unit lattice square \(K\), or (on 3D spaces) if they are \(K\)-equivalent for some unit lattice cube \(K\) in which no two diametrically opposite corners are white points that are adjacent to one another.

If we now define \textit{equivalence} between black digital loops with the same base point to be the transitive closure of immediate equivalence then Definition 2.4.1 is correct for all strongly normal spaces, in the sense that it is equivalent to Definition 3.3.4 (up to a natural group isomorphism). For a proof, see [20, Proposition 7.9.1].

4.5 Construction of the continuous analog \(C(\mathcal{P})\)

Let \(\mathcal{P}\) be a digital picture on a strongly normal space. If \(\mathcal{P}\) is two-dimensional and has black point set \(B\) then \(C(\mathcal{P})\) is given by:

\[ C(\mathcal{P}) = \cup \{ \text{co}(B \cap T) | T \subseteq T_2(\mathcal{P}) \} \]

where \text{co} means 'convex hull of', and where \(T_2(\mathcal{P})\) is the set of \((1, 1, \sqrt{2})\) triangles obtained by dividing each unit lattice square into two triangles along an appropriate diagonal, as we shall explain.

If \(\mathcal{P}\) is three-dimensional then \(C(\mathcal{P})\) is given by

\[ C(\mathcal{P}) = \cup \{ \text{co}(B' \cap T) | T \subseteq T_3(\mathcal{P}) \} \]

where \(B'\) is the \textit{augmented black point set of} \(\mathcal{P}\)—the union of the black point set of \(\mathcal{P}\) with the centroids of certain unit lattice cubes—and where \(T_3(\mathcal{P})\) is the set of tetrahedra in a certain triangulation of Euclidean 3-space. Precise definitions of \(T_3(\mathcal{P})\) and the augmented black point set will be given below.

Definition of \(T_3(\mathcal{P})\). We now define \(T_3(\mathcal{P})\) by specifying which diagonal is to be used to subdivide each unit lattice square. This definition of \(T_3(\mathcal{P})\) applies to two-dimensional and three-dimensional digital pictures \(\mathcal{P}\). When \(\mathcal{P}\) is three-dimensional we subdivide all the faces of each unit lattice cube.

Every unit lattice square must satisfy exactly one of the following three conditions:

1. The four corners are all black points or are all white points of \(\mathcal{P}\).
2. The corners are not all black points or all white points, but one of the two diagonals is a black adjacency or a white adjacency of \(\mathcal{P}\).
3. The corners are not all white points or all black points, and neither diagonal is a white adjacency or a black adjacency.

When condition 1 is satisfied, we choose the diagonal each of whose endpoints has coordinates that sum to an even number. (In fact the other diagonal would be just as suitable in this case.) When condition 2 holds, we choose the diagonal that is a black or a white adjacency. When condition 3 holds, we choose a diagonal that joins a black point to a white point (at least one diagonal has this property, by strong normality); if both diagonals join a white point to a black point then either diagonal may be used, and we again choose the diagonal each of whose endpoints has coordinates that sum to an even number.

We define \(T_3(\mathcal{P})\) to be the set of \((1, 1, \sqrt{2})\) triangles obtained by subdividing all unit lattice squares in accordance with these rules.

Definitions of \(T_3(\mathcal{P})\) and the augmented black point set. Call a unit lattice cube \(K\) special with respect to a digital picture \(\mathcal{P}\) if there are three black adjacencies and three white adjacencies in \(K\) both of which form a \((\sqrt{2}, \sqrt{2}, \sqrt{2})\) equilateral triangle, and the diameter of \(K\) perpendicular to those triangles is not a black or a white adjacency. Call a unit lattice cube ordinary if it is not special.

In a special unit lattice cube \(K\) let \(e_1, e_2, e_3, e_4, e_5\) and \(e_6\) be the edges of the two equilateral triangles formed by black and by white adjacencies (in any order), and let \(e_7\) be any straight line segment joining a corner of one triangle to the diametrically opposite corner of \(K\) (which will be on the other triangle). We subdivide \(K\) in the obvious manner into six tetrahedra whose edges are the edges of \(K\) and \(e_7\).

Every ordinary unit lattice cube is subdivided into 12 congruent tetrahedra, each of which has a corner at the centroid of the unit lattice cube and a face in \(T_3(\mathcal{P})\).

We define \(T_3(\mathcal{P})\) to be the set of tetrahedra produced by subdividing all unit lattice cubes, ordinary and special, in the ways just described.

Recall that a \textit{simple closed black digital curve in} \(\mathcal{P}\) is a connected set of black points of \(\mathcal{P}\) each of which
is adjacent to just two other points in the set. In a three-dimensional digital picture \( P = (Z^3, \beta, \omega, B) \) the augmented black point set of \( P \), denoted by \( B' \), is defined to be the union of \( B \) with the set of all centroids of ordinary unit lattice cubes \( K \) that satisfy at least one of the following two conditions:

1. One of the diameters (‘body diagonals’) of \( K \) is a black adjacency of \( P \).
2. \( K \) contains a simple closed black digital curve of \( P \) which is not contained in any one face of \( K \), and no diameter of \( K \) is a white adjacency of \( P \).

These definitions of \( I_2(P) \), \( I_3(P) \) and \( B' \) complete our definition of the continuous analog \( C(P) \) for a digital picture \( P \) on any two- or three-dimensional strongly normal space. Of course, one must prove that \( C(P) \) actually has all of the properties listed in section 4.1. This is done in [20, section 6].

5. CONCLUDING REMARKS

This paper has defined a fundamental group for digital pictures. The digital fundamental group is intuitively sound for digital picture spaces on which every digital picture has a continuous analog satisfying the conditions in section 4.1. All strongly normal spaces and spaces isomorphic to them have this property—and so do many other spaces. So the digital fundamental group is applicable to most of the digital picture spaces that have appeared in the literature on digital topology.

REFERENCES