



- (51) International Patent Classification:
G03F 1/36 (2012.01) *G03F 7/00* (2006.01)
- (21) International Application Number:
PCT/EP2024/082675
- (22) International Filing Date:
18 November 2024 (18.11.2024)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
63/610,798 15 December 2023 (15.12.2023) US
- (71) Applicant: ASML NETHERLANDS B.V. [NL/NL]; P.O. Box 324, 5500 AH Veldhoven (NL).
- (72) Inventors: ONOSE, Alexandru; P.O. Box 324, 5500 AH Veldhoven (NL). RUIZ, Hans-Christian; P.O. Box 324, 5500 AH Veldhoven (NL). BOTARI, Tiago; P.O. Box 324, 5500 AH Veldhoven (NL). VAN KRAAIJ, Markus, Gerardus, Martinus, Maria; P.O. Box 324, 5500 AH Veldhoven (NL). BOONE, Robert, Elliott; 80 W Tasman Drive, San Jose, California 95134 (US). HAMOUDA, Ayman; 80 W Tasman Drive, San Jose, California 95134 (US).

(74) Agent: ASML NETHERLANDS B.V.; Corporate Intellectual Property, P.O. Box 324, 5500 AH Veldhoven (NL).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: METHOD AND SYSTEM FOR PRIMITIVE-BASED MASK PREDICTION

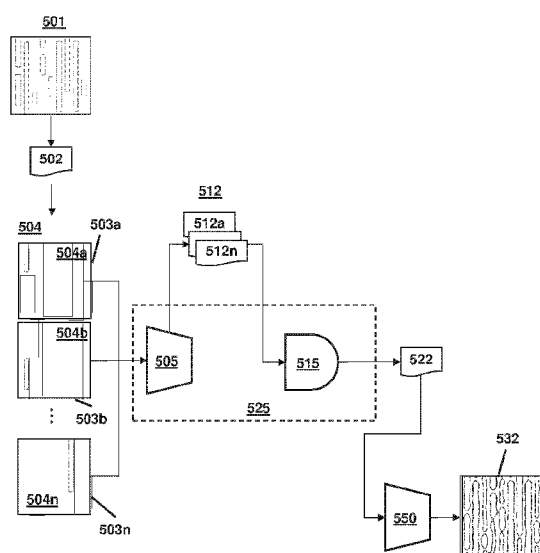


FIG. 5

(57) Abstract: Described herein is a method and system for generating a mask pattern. A set of primitive elements of a target pattern is identified. The set of primitive elements are encoded in a feature vector space to generate a set of encoded primitive elements and aggregated to generate an encoded aggregation that represents the target pattern. The encoded aggregation is input to a mask prediction model, which generates a mask image that is representative of a mask pattern for the target pattern.

Published:

- *with international search report (Art. 21(3))*
- *in black and white; the international application as filed contained color or greyscale and is available for download from PATENTSCOPE*

METHOD AND SYSTEM FOR PRIMITIVE-BASED MASK PREDICTION

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority of US application 63/610,798 which was filed on 15 December 2023 and which is incorporated herein in its entirety by reference.

TECHNICAL FIELD

[0002] The embodiments provided herein relate to semiconductor manufacturing, and more particularly to mask pattern design through computational lithography.

BACKGROUND

[0003] A lithographic apparatus is a machine that applies a desired pattern onto a target portion of a substrate. The lithographic apparatus can be used, for example, in the manufacture of integrated circuits (ICs). For example, an IC chip in a smart phone, can be as small as a person's thumbnail, and may include over 2 billion transistors. Making an IC is a complex and time-consuming process, with circuit components in different layers and including hundreds of individual steps. Errors in even one step have the potential to result in problems with the final IC and can cause device failure. High process yield and high wafer throughput can be impacted by the presence of defects.

BRIEF SUMMARY

[0004] In some embodiments, the techniques described herein relate to a method of generating a mask pattern, the method including: obtaining a set of primitive elements of a target pattern; encoding the set of primitive elements in a feature vector space to generate a set of encoded primitive elements; aggregating the set of encoded primitive elements in the feature vector space to generate an encoded aggregation that represents the target pattern; and generating a mask image based on the encoded aggregation, wherein the mask image is representative of a mask pattern for the target pattern.

[0005] In some embodiments, there is provided a non-transitory computer readable medium having instructions that, when executed by a computer, cause the computer to execute a method of any of the above embodiments.

[0006] In some embodiments, there is provided an apparatus includes a memory storing a set of instructions and a processor configured to execute the set of instructions to cause the apparatus to perform a method of any of the above embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[0007] Embodiments will now be described, by way of example only, with reference to the accompanying drawings in which:

[0008] Figure 1 illustrates a block diagram of various subsystems of a lithographic projection apparatus,

according to an embodiment.

[0009] Figure 2 is a schematic diagram of a lithographic projection apparatus, according to an embodiment.

[0010] Figure 3 illustrates an exemplary flow chart for simulating lithography in a lithographic projection apparatus, according to an embodiment.

[0011] Figures 4A and 4B are block diagrams for generating primitive elements of a target pattern, consistent with various embodiments.

[0012] Figure 5 illustrates prediction of a mask pattern from a primitive-based representation of a target pattern using prediction models, consistent with various embodiments.

[0013] Figure 6 illustrates training of a target representation model to generate an encoded aggregation of a target pattern, consistent with various embodiments.

[0014] Figure 7 illustrates training data generation for training a target representation model, consistent with various embodiments.

[0015] Figure 8 illustrates training of a mask prediction model to generate a mask pattern using an encoded aggregation of a target pattern, consistent with various embodiments.

[0016] Figure 9 illustrates training data generation for training a mask prediction model, consistent with various embodiments.

[0017] Figure 10 is a flow diagram of a method for generating a mask pattern using a set of primitive elements representing a target pattern, consistent with various embodiments.

[0018] Figure 11 is a block diagram of an example computer system, according to an embodiment.

[0019] Embodiments will now be described in detail with reference to the drawings, which are provided as illustrative examples so as to enable those skilled in the art to practice the embodiments. Notably, the figures and examples below are not meant to limit the scope to a single embodiment, but other embodiments are possible by way of interchange of some or all of the described or illustrated elements. Wherever convenient, the same reference numbers will be used throughout the drawings to refer to same or like parts. Where certain elements of these embodiments can be partially or fully implemented using known components, only those portions of such known components that are necessary for an understanding of the embodiments will be described, and detailed descriptions of other portions of such known components will be omitted so as not to obscure the description of the embodiments. In the present specification, an embodiment showing a singular component should not be considered limiting; rather, the scope is intended to encompass other embodiments including a plurality of the same component, and vice-versa, unless explicitly stated otherwise herein. Moreover, applicants do not intend for any term in the specification or claims to be ascribed an uncommon or special meaning unless explicitly set forth as such. Further, the scope encompasses present and future known equivalents to the components referred to herein by way of illustration.

DETAILED DESCRIPTION

[0020] A lithographic apparatus is a machine that applies a designed pattern onto a target portion of a substrate. This process of transferring the designed pattern to the substrate is called a patterning process. The patterning process can include a patterning step to transfer a pattern from a patterning device (such as a mask) to the substrate. Various methods are used to generate a mask pattern (e.g., continuous mask transmission (CTM)). A CTM technique is an inverse lithography solution that can generate a grayscale guidance map for a mask pattern (e.g., CTM image) for a given input design layout or a target pattern, at certain process condition values (e.g., CD, focus, dose values, etc.), source information (e.g., pupil), etc. As another example, prediction models, such as a machine learning (ML) model, are used to generate a mask pattern for a given target pattern. Typically, the ML model is trained using a training dataset of target patterns and corresponding “ground truth” mask patterns. In some embodiments, a design layout corresponds to a full-chip design layout and a target pattern is a region or a portion of the design layout. An entire mask pattern may be mapped to a target pattern as a one unit. Such conventional methods have drawbacks or limitations. For example, there is little flexibility in creating ground truth dataset, with the main task of pattern selection (e.g., selecting distinct target patterns) remaining challenging and generating ground truth for a large number of mask patterns or the entire design layout is time consuming and compute intensive. In addition, the task of constructing a model to cover an entire design space of target patterns (including patterns unknown during training) and quantifying pattern coverage is extremely challenging. Typically, the ground truth dataset includes disjoint regions from various locations of a design layout and the machine learning model solutions provided via an inverse solution of CTM may exhibit instability, producing inconsistent predicted mask patterns for the same target pattern (e.g., because the ground truth dataset provided by the CTM can be inconsistent). These and other drawbacks exist.

[0021] Disclosed are embodiments for predicting a mask pattern for a target pattern using a set of primitive elements associated with the target pattern. In some embodiments, a full-chip design layout may be represented using a collection of primitive elements, where each primitive element may be considered as a building block or a basic component that represents a portion of the design layout. In some embodiments, each target pattern in the design layout may be composed of a selected subset of the collection of primitive elements in a certain arrangement. A primitive element can be a feature, a portion of a feature or a certain arrangement of multiple features, or a combination of features that occurs in many target patterns.

[0022] According to embodiments of the present disclosure, given a set of input primitive elements of a target pattern, a target representation model is trained to generate an encoded aggregation representing the target pattern by: encoding each primitive element and its location in the target pattern in an encoding (e.g., feature vector) space to generate a set of encoded primitive elements, aggregating the set of encoded primitive elements in the encoding space to generate an encoded aggregation representing the target pattern, and decoding the encoded aggregation to generate a recovered target pattern. In this manner, the target representation model can thus map a set of primitive elements and their location (e.g.,

the encoded aggregation) to a target pattern one-to-one. In some embodiments, a mask prediction model is trained to predict a mask pattern from the encoded aggregation as generated by the target representation model. Mask patterns generated using one or more methods such as CTM can be used for training the mask prediction model. As described above, the target representation model and the mask prediction models are used in collaboration with one another to generate a mask pattern design from an arrangement of a set of primitive elements representing a target pattern. In some embodiments, the target representation model and the mask prediction model can be trained using machine learning techniques.

[0023] In some embodiments, the primitive elements associated with a design layout can be generated either based on design rules and process information to be followed in generating the design layout, or by decomposing the design layout. For example, the target representation model may be trained by using primitive elements of a full-chip design, or a target pattern, or otherwise a portion or region of the design layout, e.g., corresponding to a field of view of a lithographic apparatus.

[0024] In some embodiments, for a new design, primitives not in the training data set can be identified and used to quantify pattern coverage of the target representation model, as described in greater details below.

[0025] Such a primitive element-based prediction of a mask pattern provides various benefits. For example, by using a target pattern representation as input in the low-dimension encoding domain (e.g., encoded aggregation), the mask prediction model may be significantly less complex compared to the conventional prediction models, which use a high-dimension input such as an image of a target pattern. At least a portion of the training of this prediction model framework (e.g., training of target representation model) may be performed without the need for generating mask pattern ground truth which is typically done by a time consuming and compute-intensive method. Embodiments of the present disclosure allow for generation of fast-to-compute synthetic ground truth (e.g., synthetic target patterns and corresponding mask patterns) by performing localized perturbations in a target pattern, and updating the mask pattern in the corresponding location of the perturbation, thereby, generating more consistent ground truth mask patterns. Further, embodiments of the present disclosure provide another approach to quantify the pattern coverage/selection. For example, for a new or unseen design, the pattern coverage of the prediction models may be quantified by identifying the primitive elements of the new design that are not in the training set of primitive elements.

[0026] In the present disclosure, although specific reference may be made to the manufacture of ICs, it should be explicitly understood that the description herein has many other possible applications. For example, it may be employed in the manufacture of integrated optical systems, guidance and detection patterns for magnetic domain memories, liquid crystal display panels, thin film magnetic heads, etc. The skilled artisan will appreciate that, in the context of such alternative applications, any use of the terms “reticle”, “wafer” or “die” in this text should be considered as interchangeable with the more general terms “mask”, “substrate” and “target portion”, respectively.

[0027] In the present document, the terms “radiation” and “beam” are used to encompass all types of electromagnetic radiation, including ultraviolet radiation (e.g., with a wavelength of 365, 248, 193, 157 or 126 nm) and EUV (extreme ultra-violet radiation, e.g., having a wavelength in the range of about 5-100 nm). In the present document, the term “radiation source” or “source” is used to encompass all types of sources of radiation, including laser sources, incandescent sources, etc. which may include treatment of the radiation between the radiation source and the target or other parts of the optics, including filtering, collimating, focusing, etc.

[0028] A patterning device can comprise, or can form, one or more design layouts. The design layout can be generated utilizing CAD (computer-aided design) programs. This process is often referred to as EDA (electronic design automation). Most CAD programs follow a set of predetermined design rules in order to create functional design layouts/patterning devices. These rules are set based processing and design limitations. For example, design rules define the space tolerance between devices (such as gates, capacitors, etc.) or interconnect lines, to ensure that the devices or lines do not interact with one another in an undesirable way. One or more of the design rule limitations may be referred to as a “critical dimension” (CD). A critical dimension of a device can be defined as the smallest width of a line or hole, or the smallest space between two lines or two holes. Thus, the CD regulates the overall size and density of the designed device. One of the goals in device fabrication is to faithfully reproduce the original design intent on the substrate (via the patterning device).

[0029] The term “mask” or “patterning device” as employed in this text may be broadly interpreted as referring to a generic patterning device that can be used to endow an incoming radiation beam with a patterned cross-section, corresponding to a pattern that is to be created in a target portion of the substrate. The term “light valve” can also be used in this context. Besides the classic mask (transmissive or reflective; binary, phase-shifting, hybrid, etc.), examples of other such patterning devices include a programmable mirror array. An example of such a device is a matrix-addressable surface having a viscoelastic control layer and a reflective surface. The basic principle behind such an apparatus is that (for example) addressed areas of the reflective surface reflect incident radiation as diffracted radiation, whereas unaddressed areas reflect incident radiation as undiffracted radiation. Using an appropriate filter, the said undiffracted radiation can be filtered out of the reflected beam, leaving only the diffracted radiation behind; in this manner, the beam becomes patterned according to the addressing pattern of the matrix-addressable surface. The required matrix addressing can be performed using suitable electronic means. Examples of other such patterning devices also include a programmable LCD array. An example of such a construction is given in U.S. Patent No. 5,229,872, which is incorporated herein by reference.

[0030] The term “projection optics” as used herein should be broadly interpreted as encompassing various types of optical systems, including refractive optics, reflective optics, apertures and catadioptric optics, for example. The term “projection optics” may also include components operating according to any of these design types for directing, shaping or controlling the projection beam of radiation,

collectively or singularly. The term “projection optics” may include any optical component in the lithographic projection apparatus, no matter where the optical component is located on an optical path of the lithographic projection apparatus. Projection optics may include optical components for shaping, adjusting and/or projecting radiation from the source before the radiation passes the patterning device, and/or optical components for shaping, adjusting and/or projecting the radiation after the radiation passes the patterning device. The projection optics generally exclude the source and the patterning device.

[0031] Figure 1 illustrates a block diagram of various subsystems of a lithographic projection apparatus 10A, according to an embodiment. Major components are a radiation source 12A, which may be a deep-ultraviolet excimer laser source or other type of source including an extreme ultra violet (EUV) source (the lithographic projection apparatus itself need not have the radiation source), illumination optics which, e.g., define the partial coherence (denoted as sigma) and which may include optics 14A, 16Aa and 16Ab that shape radiation from the source 12A; a patterning device (or mask) 18A; and transmission optics 16Ac that project an image of the patterning device pattern onto a substrate plane 22A.

[0032] A pupil 20A can be included with transmission optics 16Ac. In some embodiments, there can be one or more pupils before and/or after mask 18A. As described in further detail herein, pupil 20A can provide patterning of the light that ultimately reaches substrate plane 22A. An adjustable filter or aperture at the pupil plane of the projection optics may restrict the range of beam angles that impinge on the substrate plane 22A, where the largest possible angle defines the numerical aperture of the projection optics $NA = n \sin(\Theta_{max})$, wherein n is the refractive index of the media between the substrate and the last element of the projection optics, and Θ_{max} is the largest angle of the beam exiting from the projection optics that can still impinge on the substrate plane 22A.

[0033] In a lithographic projection apparatus, a source provides illumination (i.e., radiation) to a patterning device and projection optics direct and shape the illumination, via the patterning device, onto a substrate. This is not to disclaim that the source does not itself provide patterning, directing, or shaping to the radiation or that patterning, directing, or shaping does not occur between the source and the projection optics. The projection optics may include at least some of the components 14A, 16Aa, 16Ab and 16Ac. An aerial image (AI) is the radiation intensity distribution at substrate level. A resist model can be used to calculate the resist image from the aerial image, an example of which can be found in U.S. Patent Application Publication No. US 2009-0157360, the disclosure of which is hereby incorporated by reference in its entirety. The resist model is related to properties of the resist layer (e.g., effects of chemical processes which occur during exposure, post-exposure bake (PEB) and development). Optical properties of the lithographic projection apparatus (e.g., properties of the illumination, the patterning device and the projection optics) dictate the aerial image and can be defined in an optical model. Since the patterning device used in the lithographic projection apparatus can be changed, it is desirable to separate the optical properties of the patterning device from the optical

properties of the rest of the lithographic projection apparatus including at least the source and the projection optics. Details of techniques and models used to transform a design layout into various lithographic images (e.g., an aerial image, a resist image, etc.), apply optical proximity correction (OPC) using those techniques and models and evaluate performance (e.g., in terms of process window) are described in U.S. Patent Application Publication Nos. US 2008-0301620, 2007-0050749, 2007-0031745, 2008-0309897, 2010-0162197, and 2010-0180251, the disclosure of each which is hereby incorporated by reference in its entirety.

[0034] One aspect of understanding a lithographic process is understanding the interaction of the radiation and the patterning device. The electromagnetic field of the radiation after the radiation passes the patterning device may be determined from the electromagnetic field of the radiation before the radiation reaches the patterning device and a function that characterizes the interaction. This function may be referred to as the mask transmission function (which can be used to describe the interaction by a transmissive patterning device and/or a reflective patterning device).

[0035] The mask transmission function may have a variety of different forms. One form is binary. A binary mask transmission function has either of two values (e.g., zero and a positive constant) at any given location on the patterning device. A mask transmission function in the binary form may be referred to as a binary mask. Another form is continuous. Namely, the modulus of the transmittance (or reflectance) of the patterning device is a continuous function of the location on the patterning device. The phase of the transmittance (or reflectance) may also be a continuous function of the location on the patterning device. A mask transmission function in the continuous form may be referred to as a continuous tone mask or a continuous transmission mask (CTM). For example, the CTM may be represented as a pixelated image, where each pixel may be assigned a value between 0 and 1 (e.g., 0.1, 0.2, 0.3, etc.) instead of binary value of either 0 or 1. In an embodiment, CTM may be a pixelated gray scale image, where each pixel has values (e.g., within a range [-255, 255], normalized values within a range [0, 1] or [-1, 1] or other appropriate ranges).

[0036] The thin-mask approximation, also called the Kirchhoff boundary condition, is widely used to simplify the determination of the interaction of the radiation and the patterning device. The thin-mask approximation assumes that the thickness of the structures on the patterning device is very small compared with the wavelength and that the widths of the structures on the mask are very large compared with the wavelength. Therefore, the thin-mask approximation assumes the electromagnetic field after the patterning device is the multiplication of the incident electromagnetic field with the mask transmission function. However, as lithographic processes use radiation of shorter and shorter wavelengths, and the structures on the patterning device become smaller and smaller, the assumption of the thin-mask approximation can break down. For example, interaction of the radiation with the structures (e.g., edges between the top surface and a sidewall) because of their finite thicknesses ("mask 3D effect" or "M3D") may become significant. Encompassing this scattering in the mask transmission function may enable the mask transmission function to better capture the interaction of the radiation

with the patterning device. A mask transmission function under the thin-mask approximation may be referred to as a thin-mask transmission function. A mask transmission function encompassing M3D may be referred to as a M3D mask transmission function.

[0037] Figure 2 schematically depicts an exemplary lithographic projection apparatus whose illumination source could be optimized utilizing the methods described herein. The apparatus comprises:

- an illumination system IL, to condition a beam B of radiation. In this particular case, the illumination system also comprises a radiation source SO;

- a first object table (e.g., mask table, patterning device table or reticle stage) MT provided with a patterning device holder to hold a patterning device MA (e.g., a reticle), and connected to a first positioner to accurately position the patterning device with respect to item PS;

- a second object table (substrate table or wafer stage) WT provided with a substrate holder to hold a substrate W (e.g., a resist-coated silicon wafer), and connected to a second positioner to accurately position the substrate with respect to item PS;

- a projection system ("lens") PS (e.g., a refractive, catoptric or catadioptric optical system) to image an irradiated portion of the patterning device MA onto a target portion C (e.g., comprising one or more dies) of the substrate W.

[0038] As depicted herein, the apparatus is of a transmissive type (i.e., has a transmissive mask). However, in general, it may also be of a reflective type, for example (with a reflective mask). Alternatively, the apparatus may employ another kind of patterning device as an alternative to the use of a classic mask; examples include a programmable mirror array or LCD matrix.

[0039] The source SO (e.g., a mercury lamp or excimer laser) produces a beam of radiation. This beam is fed into an illumination system (illuminator) IL, either directly or after having traversed conditioning means, such as a beam expander Ex, for example. The illuminator IL may comprise adjusting means AD for setting the outer or inner radial extent (commonly referred to as σ -outer and σ -inner, respectively) of the intensity distribution in the beam. In addition, it will generally comprise various other components, such as an integrator IN and a condenser CO. In this way, the beam B impinging on the patterning device MA has a desired uniformity and intensity distribution in its cross-section.

[0040] It should be noted with regard to Figure 2 that the source SO may be within the housing of the lithographic projection apparatus (as is often the case when the source SO is a mercury lamp, for example), but that it may also be remote from the lithographic projection apparatus, the radiation beam that it produces being led into the apparatus (e.g., with the aid of suitable directing mirrors); this latter scenario is often the case when the source SO is an excimer laser (e.g., based on KrF, ArF or F₂ lasing).

[0041] The beam B subsequently intercepts the patterning device MA, which is held on a patterning device table MT. Having traversed the patterning device MA, the beam B passes through the lens PS, which focuses the beam B onto a target portion C of the substrate W. With the aid of the second positioning means (and interferometric measuring means IF), the substrate table WT can be moved accurately, e.g., so as to position different target portions C in the path of the beam B. Similarly, the

first positioning means can be used to accurately position the patterning device MA with respect to the path of the beam B, e.g., after mechanical retrieval of the patterning device MA from a patterning device library, or during a scan. In general, movement of the object tables MT, WT will be realized with the aid of a long-stroke module (coarse positioning) and a short-stroke module (fine positioning), which are not explicitly depicted in Figure 2. However, in the case of a wafer stepper (as opposed to a step-and-scan tool) the patterning device table MT may just be connected to a short stroke actuator, or may be fixed.

[0042] The depicted tool can be used in two different modes:

- In step mode, the patterning device table MT is kept essentially stationary, and an entire patterning device image is projected in one go (i.e., a single “flash”) onto a target portion C. The substrate table WT is then shifted in the x or y directions so that a different target portion C can be irradiated by the beam B;

- In scan mode, essentially the same scenario applies, except that a given target portion C is not exposed in a single “flash”. Instead, the patterning device table MT is movable in a given direction (the so-called “scan direction”, e.g., the y direction) with a speed v, so that the projection beam B is caused to scan over a patterning device image; concurrently, the substrate table WT is simultaneously moved in the same or opposite direction at a speed $V = Mv$, in which M is the magnification of the lens PS (typically, $M = 1/4$ or $1/5$). In this manner, a relatively large target portion C can be exposed, without having to compromise on resolution.

[0043] Figure 3 illustrates an exemplary flow chart for simulating lithography in a lithographic projection apparatus, according to an embodiment. As will be appreciated, the models may represent a different patterning process and need not comprise all the models described below. A source model 300 represents optical characteristics (including radiation intensity distribution, bandwidth and/or phase distribution) of the illumination of a patterning device. The source model 300 can represent the optical characteristics of the illumination that include, but not limited to, numerical aperture settings, illumination sigma (σ) settings as well as any particular illumination shape (e.g., off-axis radiation shape such as annular, quadrupole, dipole, etc.), where σ (or sigma) is outer radial extent of the illuminator.

[0044] A projection optics model 310 represents optical characteristics (including changes to the radiation intensity distribution and/or the phase distribution caused by the projection optics) of the projection optics. The projection optics model 310 can represent the optical characteristics of the projection optics, including aberration, distortion, one or more refractive indexes, one or more physical sizes, one or more physical dimensions, etc.

[0045] The patterning device / design layout model module 320 captures how the design features are laid out in the pattern of the patterning device and may include a representation of detailed physical properties of the patterning device, as described, for example, in U.S. Patent No. 7,587,704, which is incorporated by reference in its entirety. In an embodiment, the patterning device / design layout model module 320 represents optical characteristics (including changes to the radiation intensity distribution

and/or the phase distribution caused by a given design layout) of a design layout (e.g., a device design layout corresponding to a feature of an integrated circuit, a memory, an electronic device, etc.), which is the representation of an arrangement of features on or formed by the patterning device. Since the patterning device used in the lithographic projection apparatus can be changed, it is desirable to separate the optical properties of the patterning device from the optical properties of the rest of the lithographic projection apparatus including at least the illumination and the projection optics. The objective of the simulation is often to accurately predict, for example, edge placements and CDs, which can then be compared against the device design. The device design is generally defined as the pre-OPC patterning device layout, and will be provided in a standardized digital file format such as GDSII or OASIS.

[0046] An aerial image 330 can be simulated from the source model 300, the projection optics model 310 and the patterning device / design layout model module 320. An aerial image (AI) is the radiation intensity distribution at substrate level. Optical properties of the lithographic projection apparatus (e.g., properties of the illumination, the patterning device, and the projection optics) dictate the aerial image.

[0047] A resist layer on a substrate is exposed by the aerial image and the aerial image is transferred to the resist layer as a latent “resist image” (RI) therein. The resist image (RI) can be defined as a spatial distribution of solubility of the resist in the resist layer. A resist image 350 can be simulated from the aerial image 330 using a resist model 340. The resist model can be used to calculate the resist image from the aerial image, an example of which can be found in U.S. Patent Application No. 8,200,468, the disclosure of which is hereby incorporated by reference in its entirety. The resist model 340 typically describes the effects of chemical processes which occur during resist exposure, post exposure bake (PEB) and development, in order to predict, for example, contours of resist features formed on the substrate and so it typically related only to such properties of the resist layer (e.g., effects of chemical processes which occur during exposure, post-exposure bake and development). In an embodiment, the optical properties of the resist layer, e.g., refractive index, film thickness, propagation, and polarization effects— may be captured as part of the projection optics model 310.

[0048] So, in general, the connection between the optical and the resist model is a simulated aerial image intensity within the resist layer, which arises from the projection of radiation onto the substrate, refraction at the resist interface and multiple reflections in the resist film stack. The radiation intensity distribution (aerial image intensity) is turned into a latent “resist image” by absorption of incident energy, which is further modified by diffusion processes and various loading effects. Efficient simulation methods that are fast enough for full-chip applications approximate the realistic 3-dimensional intensity distribution in the resist stack by a 3-dimensional aerial (and resist) image.

[0049] In an embodiment, the resist image 350 can be used as an input to a post-pattern transfer process model module 360. The post-pattern transfer process model module 360 defines performance of one or more post-resist development processes (e.g., etch, development, etc.).

[0050] Simulation of the patterning process can, for example, predict contours, CDs, edge placement (e.g., edge placement error), etc. in the resist and/or etched image. Thus, the objective of the simulation

is to accurately predict, for example, edge placement, and/or aerial image intensity slope, and/or CD, etc. of the printed pattern. These values can be compared against an intended design to, e.g., correct the patterning process, identify where a defect is predicted to occur, etc. The intended design is generally defined as a pre-OPC design layout which can be provided in a standardized digital file format such as GDSII or OASIS or other file format.

[0051] Thus, the model formulation describes most, if not all, of the known physics and chemistry of the overall process, and each of the model parameters desirably corresponds to a distinct physical or chemical effect. The model formulation thus sets an upper bound on how well the model can be used to simulate the overall manufacturing process.

[0052] The following paragraphs describe a system and a method for predicting a mask pattern using a set of primitive elements associated with a target pattern. A primitive element may be a feature, a portion of a feature, an arrangement of multiple features, or combination of features that occurs in many target patterns. In some embodiments, a collection of primitive elements may be used to represent an entire full-chip design layout, and a target pattern of the design layout may be represented using a subset of the collection of primitive elements. A prediction model framework (e.g., one or more prediction models such as ML models) may be used to predict a mask pattern for a given target pattern. For example, the prediction model framework may include two prediction models – a target representation model that generates an encoded aggregation from a set of primitive elements that is representative of a target pattern, and a mask prediction model that generates a mask pattern from the encoded aggregation. In some embodiments, the prediction framework may not be implemented as two separate models in which case the two prediction models may be implemented as sub-components of the prediction framework.

[0053] In the training stage, given a set of input primitive elements of a target pattern, an encoder of the target representation model is trained to encode each primitive element and its location in a target pattern in an encoding space to generate a set of encoded primitive elements, the aggregator is trained to aggregate the set of encoded primitive elements into an encoded aggregation representing the target pattern, and the decoder decodes the encoded aggregation to generate a recovered target pattern. The target representation model is trained until the difference between the recovered target pattern and the pattern is reduced (e.g., minimized). Thus, by training the target representation model to generate the encoded aggregation, it is trained to map a set of primitive elements to a target pattern one-to-one. The mask prediction model may be trained to predict a mask pattern from the encoded aggregation. In the training stage, an encoded aggregation of a target pattern (e.g., generated using the target representation model) and a ground truth mask pattern (e.g., generated using one or more known methods such as CTM method) of the target pattern is input to the mask prediction model to generate a predicted mask pattern. For example, the mask prediction model is trained until the difference between the predicted mask pattern and the ground truth mask pattern is reduced (e.g., minimized).

[0054] Figures 4A and 4B are block diagrams for generating primitive elements of a target pattern, consistent with various embodiments. A design layout contains multiple features to be printed on a substrate. In some embodiments, the design layout may be represented as a collection of primitive elements, where each primitive element may be considered as a building block or a basic component that represents a portion of the design layout. In some embodiments, each target pattern in the design layout may be composed of a selected subset of the collection of primitive elements in a certain arrangement. A primitive element can be a feature, a portion of a feature or a combination of features that occurs in many target patterns. For example, in a target pattern 404 there are many features that occur in multiple target patterns of a design layout of which the target pattern 404 is a part. As an example, a rectangular feature 405a with a first set of dimensions may occur in multiple target patterns of the design layout. Accordingly, the rectangular feature 405a with the first set of dimensions may be selected as a first primitive element 405a, which may be used to represent each of the three occurrences (a first occurrence 411a, a second occurrence 411b and a third occurrence 411c) of the feature in the target pattern 404 (e.g., and other target patterns of the design layout). Similarly, other features that occur in multiple target patterns of the design layout may be selected as primitive elements, such as primitive elements 405b, 405c, 405d, and 405e, which may be used to represent any occurrences of those features in the target pattern 404.

[0055] In some embodiments, a combination of features that occurs in multiple target patterns of the design layout may be selected as a primitive element. For example, a second primitive element 405f represents a combination of two rectangular features, a third primitive element 405g represents a combination of two circular features, and a fourth primitive element 405h represents a combination of two circular features and a rectangular feature each of which occurs in multiple target patterns. The primitive elements may be used to represent an entire full chip design layout, or a target pattern of the design layout.

[0056] In some embodiments, if a feature is unique in the design layout, that is, appears not more than once in the design layout, then either the feature is considered as a primitive or a portion of that feature that occurs in multiple target patterns may be selected as a primitive element in which case the occurrence of such a feature in a target pattern can be represented using two or more primitive elements.

[0057] In some embodiments, a collection of primitive elements 405 includes the primitive elements that can be used to represent an entire full-chip design layout of which the target pattern 404 is a part. The target pattern 404 may be represented using a selected subset of the collection of primitive elements 405 along with the location of their occurrence in the target pattern 404. An example representation of the target pattern using the primitive elements may be – ((first primitive element 405a, location of first occurrence 411a), (first primitive element 405a, location of second occurrence 411b)... (second primitive element 405b, location of first occurrence), (second primitive element 405b, location of second occurrence)... (n^{th} primitive element, location of occurrence) and so on). In some embodiments, a location of the feature may be a location in a region of the target pattern 404 where the region

corresponds to a field of view of the lithographic apparatus. In some embodiments, the location may be represented using coordinates (e.g., (x, y) coordinates). In some embodiments, the location may correspond to the location of the primitive element in an image of the primitive element or the target pattern 404, which may be derived from the corresponding image. Additional details with respect to the location are described at least with reference to Figure 5 below.

[0058] The primitive elements of a given design layout can be generated either by decomposing the given design layout (e.g., the design layout of which the target pattern 404 is a part), or based on design rules (e.g., design rules 402). For example, as illustrated in Figure 4A, the set of primitive elements 405 may be generated by decomposing the design layout to its primitive elements. In some embodiments, decomposing the design layout includes analyzing or processing the design layout (e.g., using pattern matching, heuristic matching, image segmentation or other known feature extraction methods) to identify those of the features, portions of features or combinations of features that repeatedly occurs in multiple target patterns of the design layout as the primitive elements. After obtaining the collection of primitive elements 405, a set of primitive elements that are associated with the target pattern 404 may be identified. For example, the target pattern 404 may be compared with the collection of primitive elements 405 (e.g., using pattern matching, heuristic matching, image segmentation or other known feature extraction methods) to identify those of the primitive elements 405 that match with a feature, portion of a feature, or combination of features in the target pattern 501 as a set of primitive elements that are associated with the target pattern 404. The identified primitive elements along with the location of their occurrences in the target pattern 404 may be used to represent the target pattern 404 (e.g., as described above).

[0059] As illustrated in Figure 4B, the primitive elements may be generated using design rules 402. In some embodiments, the design rules 402 may include information regarding design rules or process information to be followed in generating a design layout. For example, the design rules may indicate a pitch of the features (e.g., distance between two features in x or y direction), critical dimension (CD) of a feature, minimum CD, maximum CD, or other such attributes of the features, information regarding a lithographic apparatus, or information regarding a lithographic process to be followed for the generation of or printing of the design layout on a substrate. In the first example 490 of design rules for a first design layout, the design rules indicate a pitch 451 in x direction and a pitch 452 in y direction, and a CD 453 of a feature (e.g., width and height) based on which a first target pattern 450 of the first design layout may be generated. The first target pattern 450 may be decomposed to identify the primitive elements. For example, the square shaped feature with the same dimension may be considered to occur in multiple target patterns of the first design layout, and therefore, the square shaped feature may be identified as a primitive element 455, and this primitive element 455 may be used to represent the first target pattern 450. That is, the entire target pattern 450 may be represented using the primitive element 455 and the location of each occurrence of the primitive element 455 in the target pattern 450. In another example, a horizontal pair of square features (e.g., one next to the other in x direction) may

be considered to occur across multiple target patterns, and therefore, the combination of the two horizontal square features may be identified as a single primitive element 456. In yet another example, a vertical pair of square features (e.g., one below/above the other in y direction) may be considered to occur across multiple target patterns, and therefore, the combination of the two vertical square features may be identified as a single primitive element 457.

[0060] In the second example 492 of design rules for a second design layout, the design rules indicate a pitch 461 in x direction, and a CD 462 of a feature (e.g., width) based on which a second target pattern 470 of the second design layout may be generated. The second target pattern 470 may be decomposed to identify the primitive elements. For example, a square shaped feature with the same dimension, a first rectangle feature with a first set of dimensions, and a second rectangle feature with a second set of dimensions may be considered to occur across multiple target patterns of the second design layout, and each of the square shaped feature, the first rectangle feature and the second rectangle feature may be identified as primitive element 465a, primitive element 465b, and primitive element 465c, respectively. That is, the entire second target pattern 470 may be represented using just the three primitive elements 465a-465c. In another example, a primitive element may be a parameterized primitive element. In some embodiments, a parameterized primitive element is associated with a parameter that is descriptive of an attribute of the feature and the parameter has different values for different occurrences of the feature. For example, the second target pattern 470 may be represented using a single parameterized primitive element 475 having a height of the feature as a parameter 476. The square shaped features may be represented using the parameterized primitive element 475 having a first value for the parameter 476 (e.g., the height of the square), a first rectangle feature of a first height may be represented using the parameterized primitive element 475 having a second value for the parameter 476, and a second rectangle feature of a second height may be represented using the parameterized primitive element 475 having a third value for the parameter 476.

[0061] While the above examples illustrate polygonal features, the features are not limited to the polygonal features and may have curvilinear features. The curvilinear features may also be represented using primitive elements according to embodiments of the present disclosure.

[0062] The primitive elements may be represented in a number of formats. In some embodiments, a primitive element may be represented as an image. In some embodiments, a primitive element may be represented using a set of coordinates or a set of attributes defining the corresponding primitive element. For example, a rectangular primitive element may be represented using a starting coordinate of the feature (e.g., (x,y) coordinates of a vertex of the feature) and a width and height of the feature. In another example, a rectangular primitive element may be represented using coordinates of diagonally opposite vertices of the feature (e.g., (x1,y1) and (x2,y2) coordinates of diagonally opposite vertices of the feature) in which case the coordinates of the remaining vertices may be derived from the provided coordinates. In another example, a rectangular or other polygonal primitive element may be represented using coordinates of all vertices of the feature. In another example, a circular primitive element may

be represented using coordinates of a center of the circle and a radius of the circle. In another example, a curvilinear primitive element may be represented using (a) coordinates of a set of points forming the curvilinear feature, and (b) information regarding a tangent or directional derivative information for each point of the set of points. In another example, a primitive element may be represented as a set of elements (e.g., at least one of vertices, edges, faces, or polygons) defining the corresponding primitive element.

[0063] A mask pattern corresponding to a target pattern may be predicted, or generated, using primitive-based representation of the target pattern. The following paragraphs describe prediction of the mask pattern from a primitive-based representation of a target pattern using a prediction framework.

[0064] Figure 5 illustrates prediction of a mask pattern from a primitive-based representation of a design layout using prediction models, consistent with various embodiments. A set of primitive elements 502 that represents a target pattern 501 (e.g., a region of an entire full-chip design layout) along with their location in the target pattern 501 is provided as a set of inputs to a prediction model, such as a target representation model 525, which is trained to generate an encoded aggregation 522 that is representative of the target pattern 501. The encoded aggregation 522 is provided as an input to a second prediction model, such as a mask prediction model 550, which is trained to generate a mask pattern 532 (e.g., as a mask image) for the target pattern 501.

[0065] The set of primitive elements 502 that represents the target pattern 501 is identified. The set of primitive elements 502 may be identified in a number of ways. For example, the target pattern 501 may be compared with a collection of primitive elements (e.g., using pattern matching, heuristic matching, image segmentation or other known feature extraction methods) representing one or more design layouts (e.g., design layout of which the target pattern 501 is a part) to identify those of the primitive elements that match with a feature, portion of a feature, or combination of features in the target pattern 501 as the set of primitive elements 502. The collection of primitive elements may be generated using one or more methods (e.g., by decomposing one or more design layouts, such as the design layout of which the target pattern 501 is a part, or based on design rules, as described at least with reference to Figures 4A and 4B above). Further, in some embodiments, the collection of primitive elements may be those of the primitive elements that are used in training of the target representation model 525 (which is described below in further detail).

[0066] The set of primitive elements 502 are provided as a set of inputs 504 to the target representation model 525. Each input includes information regarding a primitive element and a location of the primitive element in the target pattern 501. For example, a first input 504a may include (a) a first primitive element that is representative of a first feature in the target pattern 501 and (b) a location of a first occurrence of the first primitive element in the target pattern 501 (e.g., in a particular area of the target pattern 501 corresponding to the field of view), a second input 504b may include (a) the first primitive element and (b) a location of the second occurrence of the first primitive element, and a third input 504c may include (a) a second primitive element that is representative of a second feature and (b)

a location of the first occurrence of the second primitive element. Such inputs may be provided for all occurrences of all features in the target pattern 501 using the set of primitive elements 502.

[0067] The location of a primitive element may be specified in different ways. For example, if a primitive element is represented as an image (e.g., as illustrated by images 503a-503n), then the location of the primitive element may be inherent in the image and may be derived from the image. For example, in the first input 504a, the primitive element is located in the first location of the image 503a, and the first location may be derived from the image 503a. Similarly, in the second input 504b, the primitive element is located in a second location of the image 503b, and the second location may be derived from the image 503b. In another example, if the primitive elements are represented as a set of coordinates, a set of points, polygons, etc., the location information may be included as location coordinates in the representation. Accordingly, the set of inputs 504a-504n, which includes the set of primitive elements 502 and their locations, together represent the target pattern 501.

[0068] An encoder 505 of the target representation model 525 receives the set of primitive elements 502 and their locations as the set of inputs 504a-504n. The encoder 505 encodes a primitive element and its location to generate an encoded primitive element in an encoding space (e.g., feature vector space). For example, the encoder 505 encodes a first input 504a having a first primitive element and its location in the target pattern 501 to generate a first encoded primitive element 512a. Similarly, the encoder 505 encodes each input of the set of inputs 504 (e.g., 504a-504n) to generate a set of encoded primitive elements 512 (e.g., 512a-512n). In some embodiments, the encoder 505 reduces the input dimensions and compresses the input data into an encoded representation (e.g., from high-dimension to low-dimension). The encoding function may be represented as:

$$f(P_i) \rightarrow e_i$$

... Eq. (1)

where $f(.)$ represents the encoder 505, P_i represents i^{th} input (e.g., primitive element and its location), and e_i represents i^{th} encoded primitive element.

[0069] An aggregator 515 of the target representation model 525 is trained to aggregate the set of encoded primitive elements 512a-512n to generate an encoded aggregation 522. The encoded aggregation is representative of the target pattern 501. The aggregation may be performed in a number of ways. For example, the aggregation may be an aggregation of the set of encoded primitive elements in the encoding space (e.g., feature vector space). In some embodiments, the aggregation may be represented as:

$$g(e_1, \dots, e_n) \rightarrow a$$

... Eq. (2)

where $g(.)$ represents aggregation operation of the aggregator 515, and a represents the encoded aggregation.

[0070] The encoded aggregation 522 is provided as input to the mask prediction model 550, which is trained to predict a mask pattern for any given target pattern. The mask prediction model 550 generates the mask pattern 532 based on the encoded aggregation 522.

[0071] The prediction models described at least with reference to Figure 5 are trained to generate their corresponding outputs. For example, the target representation model is trained using a training dataset having several target patterns of one or more design layouts as ground truth and primitive elements for each of the several ground truth target patterns to generate an encoded aggregation. Similarly, the mask prediction model is trained using a training dataset having encoded aggregations for several target patterns and ground truth mask patterns corresponding to the target patterns to generate a mask pattern. The following paragraphs describe training of the target representation model and the mask prediction model.

[0072] Figure 6 illustrates training of a target representation model to generate an encoded aggregation of a target pattern, consistent with various embodiments. The training dataset used to train the target representation model 525 may include (a) several training target patterns of one or more design layouts as ground truth (also referred to as “training target pattern”), and (b) several sets of primitive elements where each set of primitive elements (also referred to as “training set of primitive elements”) represents a corresponding target pattern. For example, the training dataset may include a target pattern 602 and a set of primitive elements 603 representing the target pattern 602. In some embodiments, the target pattern 602 corresponds to a portion or a region of an entire full-chip design layout.

[0073] The set of primitive elements 603 may be identified from a collection of primitive elements that represents a design layout of which the target pattern 602 is a part (e.g., as described at least with reference to Figures 4A and 5 above). The set of primitive elements 603 may be provided as a set of inputs 604 to the target representation model 525. As mentioned above at least with reference to Figure 5, each input includes information regarding a primitive element corresponding to a feature of the target pattern 602 and a location of the primitive element in the target pattern 602. For example, a first input 604a may include (a) a first primitive element that is representative of a first feature in the target pattern 602 and (b) a location of a first occurrence of the first primitive element in the target pattern 602, a second input 604b may include (a) the first primitive element that is representative of the first feature and (b) a location of the second occurrence of the first primitive element, and so on for all occurrences of all features in the target pattern 602. Accordingly, the set of inputs 604a-604n, which includes the set of primitive elements 603 and their locations, together represent the training target pattern 602.

[0074] The encoder 505 of the target representation model 525 receives the set of primitive elements and their locations as the set of inputs 604 and encodes the primitive elements and their location to generate a set of encoded primitive elements (also referred to as “training set of encoded primitive

elements”) in an encoding space (e.g., feature vector space). For example, the encoder 505 encodes a first input 604a having a first primitive element and its location to generate a first encoded primitive element 614a. Similarly, the encoder 505 encodes the remaining inputs to generate a set of encoded primitive elements 614a-614n.

5 [0075] An aggregator 515 of the target representation model 525 aggregates the set of encoded primitive elements 614a-614n to generate an encoded aggregation 624 (also referred to as “training encoded aggregation”). The encoded aggregation 624 is representative of the target pattern 602.

[0076] The training encoded aggregation 624 is input to a decoder 605 of the target representation model 525, which decodes the training encoded aggregation 624 to a reconstructed target pattern 652.

10 The reconstructed target pattern 652 may be compared with the target pattern 602 to compute a cost function that is indicative of a difference between the two target patterns. For example, the cost function may be represented as:

$$\arg\min_{f,d} h\left(d\left(\sum_i f\left(P_i^{(s)}\right)\right), T^{(s)}\right)$$

15 ... Eq. (3)

where $h(\cdot)$ is a data fidelity metric across all training samples s , e.g., (mean squared error (MSE)), T is a target pattern, $d(\cdot)$ is a decoding function.

[0077] The parameters of the target representation model 525 (e.g., weights or biases of any of the components – encoder 505, aggregator 515 or decoder 605) are adjusted to reduce the cost function, and the target representation model 525 is executed/trained again with the training dataset (e.g., same target pattern and same set of primitive elements, or other target patterns and corresponding set of primitive elements). The training process may be an iterative process in which each iteration includes (i) executing the target representation model 525 by inputting (a) a set of primitive elements representing a target pattern and (b) the target pattern as the ground truth, (ii) generating an encoded aggregation representing the target pattern, (iii) decoding the encoded aggregation to recover a reconstructed target pattern, (iv) computing the cost function, and (v) adjusting the parameters of the target representation model 525 to reduce the cost function. The iterations may be continued until a specified training condition is satisfied. For example, the training process may be continued for a specified number of iterations. In another example, the training process may be continued until the cost function, that is, a difference between the reconstructed target pattern (e.g., reconstructed target pattern 652) and the ground truth target pattern (e.g., target pattern 602) is reduced (e.g., minimized). Once the training condition is satisfied, the target representation model 525 is considered to be trained, and the target representation model 525 may be deployed to generate an encoded aggregation for any new target

pattern (e.g., unseen target pattern or a target pattern that the target representation model 525 is neither trained on, nor executed on to generate the encoded aggregation), as illustrated in Figure 5 above.

[0078] The encoder 505 and decoder 605 used in the target representation model 525 can be implemented using any deep learning techniques. Typically, an encoder is a neural network architecture that is used for tasks such as dimensionality reduction, feature extraction and data compression. The main function of an encoder is to map the input data (e.g., a high dimensional input such as an image of a primitive element) to a lower-dimensional representation, called the encoded space (e.g., set of encoded primitive elements). A decoder is a neural network architecture that is typically used in conjunction with an encoder to map the encoding space (e.g., aggregated encoded primitive elements) back to the original data space (e.g., reconstructed image of primitive element). In some embodiments, by first compressing the input data, the inference problem may be significantly simplified compared to performing regression on the high dimensional input directly. Such an encoder-decoder model may be implemented using any of a number of deep learning techniques (e.g., autoencoder).

[0079] Figure 7 illustrates training data generation for training a target representation model, consistent with various embodiments. To aid the training of the target representation model 525, several ground truth target patterns may be synthetically generated in a fast and efficient manner. In some embodiments, a set of primitive elements representing a given target pattern may be identified (e.g., as described at least with reference to Figure 5). Given a set of primitive elements, a number of random target patterns may be generated using the set of primitive elements. For example, given the set of primitive elements 405, a number of additional target patterns, such as a first target pattern 706 and a second target pattern 708, may be generated in a random manner (e.g., by placing different primitive elements at different locations randomly). While the target patterns may be generated in a random manner, the target pattern generation may still have to adhere to, or follow, the lithographic process design rules prescribed for generating a design layout to be printed on the substrate. Such random generation of target patterns aids in fast and efficient generation of training data for training the target representation model 525.

[0080] Figure 8 illustrates training of a mask prediction model to generate a mask pattern using an encoded aggregation of a target pattern, consistent with various embodiments. The training dataset used to train the mask prediction model 550 may include (a) several encoded aggregations each of which represents a target pattern and is generated using a trained target representation model 525 (e.g., as described at least with reference to Figure 5), and (b) several mask patterns corresponding to the target patterns as ground truth. For example, the training dataset may include (a) an encoded aggregation 824 representing a target pattern 804, and (b) a mask pattern 806 corresponding to the target pattern 804 as ground truth. The encoded aggregation 824 may be generated using a trained target representation model 525 (e.g., using a set of primitive elements corresponding to the target pattern 804 as described at least with reference to Figure 5). The ground truth mask pattern 806 may be generated using any of a number of known methods (e.g., CTM process).

[0081] The encoded aggregation 824 is input to the mask prediction model 550, which generates a predicted mask pattern 826. The predicted mask pattern 826 is compared with the ground truth mask pattern 806 to compute a cost function that is indicative of a difference between the predicted mask pattern 826 and the ground truth mask pattern 806. For example, the cost function may be represented as:

$$\arg \min_p h(p(a^{(s)}), M^{(s)})$$

... Eq. (3)

where $h(\cdot)$ is a data fidelity metric across all training samples s , e.g., MSE, M is a mask pattern, $p(\cdot)$ is the mask prediction model.

[0082] The parameters of the mask prediction model 550 (e.g., weights or biases) may be adjusted to reduce the cost function, and the mask prediction model 550 may be executed/trained again with the training dataset (e.g., the same encoded aggregation 824 and the same ground truth mask pattern 806, or other encoded aggregations corresponding to other target patterns and mask patterns corresponding to the other target patterns). The training process may be an iterative process in which each iteration includes (i) executing the mask prediction model 550 by inputting (a) an encoded aggregation representing a target pattern and (b) the mask pattern corresponding to the target pattern as the ground truth, (ii) generating a predicted mask pattern, (iii) computing the cost function, and (iv) adjusting the parameters of the target representation model 525 to reduce the cost function. The iterations may be continued until a specified training condition is satisfied. For example, the training process may be continued for a specified number of iterations. In another example, the training process may be continued until the cost function, that is, a difference between the predicted mask pattern (e.g., predicted mask pattern 826) and the ground truth mask pattern (e.g., ground truth mask pattern 806) is reduced (e.g., minimized). Once the training condition is satisfied, the mask prediction model 550 is considered to be trained, and the mask prediction model 550 may be deployed to generate or predict a mask pattern for any new target pattern (e.g., unseen target pattern or a target pattern whose encoded aggregation the mask prediction model 550 is neither trained on, nor executed on), as illustrated in Figure 5 above.

[0083] In some embodiments, the mask prediction model 550 can be trained using representative target patterns, which can be selected based on the corresponding encoded aggregations generated by the target representation model 525. For example, one of the multiple target patterns may be selected as a representative pattern based on a distance metric, e.g., a distance between a set of aggregated encodings of the target patterns in the encoding space. For example, a first target pattern corresponding to a first encoded aggregation whose distance with other encoded aggregations is above a specified threshold may be selected as a representative pattern, and the first encoded aggregation along with a first mask pattern corresponding to the first target pattern may be used for training the mask prediction model 550.

[0084] Figure 9 illustrates training data generation for training a mask prediction model, consistent with various embodiments. To aid the training of the mask prediction model 550, several ground truth mask patterns may be generated in a fast and efficient manner. A new target pattern-mask pattern ground truth pair may be generated by performing localized changes to given target pattern-mask pattern pair. In some embodiments, a new target pattern-mask pattern ground truth pair may be generated by perturbing one or more primitive elements in a given target pattern (e.g., changing location of a primitive element) and updating the corresponding mask pattern in a location corresponding to the one or more primitive elements (e.g., instead of generating an entirely new mask pattern, which is a compute-intensive process), thereby reducing an amount of time and computing resources required for generating the mask pattern. For example, given a target pattern 804-mask pattern 806 pair, a new target pattern 914 may be generated by perturbing a primitive element from the target pattern 804 (e.g., changing the location of the primitive element from a first location 908 to a second location 918) to generate a new target pattern 914. The mask pattern 806 may be updated locally - in an area 912 corresponding to the location 918 of the perturbed primitive element - (e.g., using known methods) to obtain an updated mask pattern 916 corresponding to the new target pattern 914. In some embodiments, since the location where the perturbation occurs is known, the mask pattern updating method may be restricted, or configured, to update the mask pattern locally (e.g., in an area corresponding to the location of the perturbed primitive element). Thus, by generating a new mask pattern from performing localized updates on an existing mask pattern, an amount of time and computing resources that may have otherwise been required for generating an entirely new mask pattern (e.g., using CTM process) is reduced significantly. In some embodiments, after generating a new target pattern-mask pattern pair, an encoded aggregation for the new target pattern may be generated (e.g., using a trained target representation model 525 as described at least with reference to Figure 5), which may then be used along with new mask pattern as training data for training the mask prediction model 550. Several such new target pattern-mask pattern pairs may be generated by perturbing one or more of several primitive elements of the target pattern, and the encoded aggregations generated for the new target patterns (e.g., using the trained target representation model 525) may then be used along with the new mask patterns as training data for training the mask prediction model 550.

[0085] In some embodiments, the above primitive element-based prediction of a mask pattern provides various flexibilities in training of the target representation model 525 and the mask prediction model 550. The target representation model 525 and the mask prediction model 550 may be trained in a number of ways. For example, both models may be trained using regions selected from a given design layout. In another example, the target representation model 525 may be trained using regions from the given design layout and synthetically generated target patterns (e.g., described at least with reference to Figure 7), whereas the mask prediction model 550 may be trained only using regions selected from the given design layout and the associated ground truth mask pattern. In yet another example, both the models may be trained using regions from the given design layout and synthetically generated target

patterns. Another benefit of the above primitive element-based prediction of a mask pattern is that the target representation model and the mask prediction model may be trained simultaneously (or sequentially), which provides the benefit of generating ground truth mask pattern in parallel to the training of the target representation model, thereby saving time consumed in training the models.

5 Further, a portion of the training of the prediction models, e.g., training of the target representation model 525, does not even need a mask pattern, that is, the training of the target representation model 525 may be performed without the need for a mask pattern.

[0086] Figure 10 is a flow diagram of a method for generating a mask pattern using a set of primitive elements representing a target pattern, consistent with various embodiments. The method of Figure 10
10 is described at least with reference to Figure 5 above.

[0087] At process P1002, a set of primitive elements of a target pattern is obtained. For example, the set of primitive elements 502 for a target pattern 501 is obtained. In some embodiments, a design layout may be represented as a collection of primitive elements, where each primitive element may be considered as a building block or a basic component that represents a portion of the design layout. In
15 some embodiments, each target pattern in the design layout may be composed of selected subset of the collection of primitive elements in a certain arrangement. A primitive element can be a feature, a portion of a feature or a combination of features that occurs in many target patterns. The set of primitive elements 502 may be identified from a collection of primitive elements that represent a design layout of which the target pattern 501 is a part (e.g., as described at least with reference to Figure 5 above).

[0088] At process P1004, the set of primitive elements representing the target pattern is encoded in a feature vector space to generate a set of encoded primitive elements. In some embodiments, the encoding process may encode the set of primitive elements along with their locations in the target pattern. For example, the set of primitive elements 502 along with their locations in the target pattern 501 are provided as a set of inputs 504a-504n to an encoding process, which encodes the set of inputs
20 504a-504n to generate a corresponding set of encoded primitive elements 512a-512n. In some embodiments, the encoding process reduces the input dimensions and compress the input data into an encoded representation (e.g., from high-dimension to low-dimension). In some embodiments, the encoding process may be performed using a prediction model, such as an encoder 505 of a target representation model 525, as described at least with reference to Figure 5.

[0089] At process P1006, the set of encoded primitive elements are aggregated to generate an encoded aggregation. For example, the set of encoded primitive elements 512a-512n are aggregated in the feature vector space to generate an encoded aggregation 522. The encoded aggregation 522, which is an aggregation of the set of encoded primitive elements 512a-512n, is representative of the target pattern 501. The aggregation process may perform the aggregation in a number of ways (e.g., using known
25 methods). In some embodiments, the aggregation process is performed by an aggregator 515 of the target representation model 525, as described at least with reference to Figure 5. In some embodiments,

the target representation model 525 is trained to generate an encoded aggregation of the primitive elements that is representative of a target pattern, as described at least with reference to Figure 6.

[0090] At process P1008, a mask image having a mask pattern corresponding to the target pattern is generated based on the encoded aggregation. For example, the encoded aggregation 522 is input to a mask prediction model 550, which generates the mask pattern 532. In some embodiments, the mask prediction model 550 is trained to generate a mask pattern for a given target pattern from an encoded aggregation of the given target pattern, as described at least with reference to Figure 8 above.

[0091] In some embodiments, such a primitive element-based prediction of a mask pattern provides various benefits. For example, the mask prediction model 550 may be significantly less complex compared to the conventional prediction models because the input (e.g., encoded aggregation) is in the low-dimension encoding domain. As another example, at least a portion of the training of the prediction models (e.g., training of target representation model) may be performed without the need for mask pattern ground truth. As another example, the method allows for generation of fast-to-compute synthetic ground truth (e.g., synthetic target patterns and corresponding mask patterns described at least with reference to Figures 7 and 9) by performing localized perturbations in a target pattern, and updating the mask pattern in the corresponding location of the perturbation, thereby, generating more consistent ground truth mask patterns.

[0092] As yet another example of the benefits, the above methods may be used to determine or quantify pattern coverage or selection of the prediction models. In some embodiments, a prediction accuracy of the mask prediction model 550 may be determined based on the pattern coverage. The pattern coverage may be determined in a number of ways. For example, the pattern coverage is indicative of whether an input target pattern for which a mask pattern has to be predicted using the mask prediction model 550 can be represented using the set of primitive elements that was used in training of the target representation model 525. The prediction accuracy may be considered to be above a specified accuracy threshold if the input target pattern can be represented using the set of primitive elements used in training the target representation model 525. That is, the prediction accuracy may be considered to be above a specified accuracy threshold if the primitive elements used to represent the input target pattern is part of the set of primitive elements used in training the target representation model 525. As another example, the pattern coverage may be indicative of a difference between the input target pattern and a reconstructed input target pattern generated by a decoder 605 of the target representation model 525. The prediction accuracy may be considered to be above a specified accuracy threshold based on the difference between the two target patterns being below a specified threshold. As another example, the pattern coverage may be indicative of a probability that an encoded aggregation generated by the target representation model 525 for the input target pattern belongs to a distribution of the several encoded aggregations (e.g., in a feature vector space) used in training the mask prediction model 550. In some embodiments, the probability of the encoded aggregation belonging to the distribution of the several encoded aggregations used in training the mask prediction model 550 may be determined in a number

of ways. For example, the probability may be determined using a similarity metric, which is determined based on a distance between the encoded aggregation and one or more encoded aggregations (e.g., in a feature vector space) of the several encoded aggregations used in training the mask prediction model 550. The prediction accuracy may be considered to be above a specified accuracy threshold based on the similarity metric being above a specified threshold. Thus, the prediction accuracy aids in quantifying the pattern coverage of the prediction models.

[0093] Figure 11 is a block diagram that illustrates a computer system 100 which can assist in implementing various methods and systems disclosed herein. The computer system 100 may be used to implement any of the entities, components, modules, or services depicted in the examples of the figures (and any other entities, components, modules, or services described in this specification). The computer system 100 may be programmed to execute computer program instructions to perform functions, methods, flows, or services (e.g., of any of the entities, components, or modules) described herein. The computer system 100 may be programmed to execute computer program instructions by at least one of software, hardware, or firmware.

[0094] Computer system 100 includes a bus 102 or other communication mechanism for communicating information, and a processor 104 (or multiple processors 104 and 105) coupled with bus 102 for processing information. Computer system 100 also includes a main memory 106, such as a random access memory (RAM) or other dynamic storage device, coupled to bus 102 for storing information and instructions to be executed by processor 104. Main memory 106 also may be used for storing temporary variables or other intermediate information during execution of instructions to be executed by processor 104. Computer system 100 further includes a read only memory (ROM) 108 or other static storage device coupled to bus 102 for storing static information and instructions for processor 104. A storage device 110, such as a magnetic disk or optical disk, is provided and coupled to bus 102 for storing information and instructions.

[0095] Computer system 100 may be coupled via bus 102 to a display 112, such as a cathode ray tube (CRT) or flat panel or touch panel display for displaying information to a computer user. An input device 114, including alphanumeric and other keys, is coupled to bus 102 for communicating information and command selections to processor 104. Another type of user input device is cursor control 116, such as a mouse, a trackball, or cursor direction keys for communicating direction information and command selections to processor 104 and for controlling cursor movement on display 112. This input device typically has two degrees of freedom in two axes, a first axis (e.g., x) and a second axis (e.g., y), that allows the device to specify positions in a plane. A touch panel (screen) display may also be used as an input device.

[0096] According to one embodiment, portions of one or more methods described herein may be performed by computer system 100 in response to processor 104 executing one or more sequences of one or more instructions contained in main memory 106. Such instructions may be read into main memory 106 from another computer-readable medium, such as storage device 110. Execution of the

sequences of instructions contained in main memory 106 causes processor 104 to perform the process steps described herein. One or more processors in a multi-processing arrangement may also be employed to execute the sequences of instructions contained in main memory 106. In an alternative embodiment, hard-wired circuitry may be used in place of or in combination with software instructions.

5 Thus, the description herein is not limited to any specific combination of hardware circuitry and software.

[0097] The term “computer-readable medium” as used herein refers to any medium that participates in providing instructions to processor 104 for execution. Such a medium may take many forms, including but not limited to, non-volatile media, volatile media, and transmission media. Non-volatile media
10 include, for example, optical or magnetic disks, such as storage device 110. Volatile media include dynamic memory, such as main memory 106. Transmission media include coaxial cables, copper wire and fiber optics, including the wires that comprise bus 102. Transmission media can also take the form of acoustic or light waves, such as those generated during radio frequency (RF) and infrared (IR) data communications. Common forms of computer-readable media include, for example, a floppy disk, a
15 flexible disk, hard disk, magnetic tape, any other magnetic medium, a CD-ROM, DVD, any other optical medium, punch cards, paper tape, any other physical medium with patterns of holes, a RAM, a PROM, and EPROM, a FLASH-EPROM, any other memory chip or cartridge, a carrier wave as described hereinafter, or any other medium from which a computer can read.

[0098] Various forms of computer readable media may be involved in carrying one or more sequences
20 of one or more instructions to processor 104 for execution. For example, the instructions may initially be borne on a magnetic disk of a remote computer. The remote computer can load the instructions into its dynamic memory and send the instructions over a telephone line using a modem. A modem local to computer system 100 can receive the data on the telephone line and use an infrared transmitter to convert the data to an infrared signal. An infrared detector coupled to bus 102 can receive the data carried in the
25 infrared signal and place the data on bus 102. Bus 102 carries the data to main memory 106, from which processor 104 retrieves and executes the instructions. The instructions received by main memory 106 may optionally be stored on storage device 110 either before or after execution by processor 104.

[0099] Computer system 100 also preferably includes a communication interface 118 coupled to bus
30 102. Communication interface 118 provides a two-way data communication coupling to a network link 120 that is connected to a local network 122. For example, communication interface 118 may be an integrated services digital network (ISDN) card or a modem to provide a data communication connection to a corresponding type of telephone line. As another example, communication interface 118 may be a local area network (LAN) card to provide a data communication connection to a compatible LAN. Wireless links may also be implemented. In any such implementation,
35 communication interface 118 sends and receives electrical, electromagnetic, or optical signals that carry digital data streams representing various types of information.

[0100] Network link 120 typically provides data communication through one or more networks to other data devices. For example, network link 120 may provide a connection through local network 122 to a host computer 124 or to data equipment operated by an Internet Service Provider (ISP) 126. ISP 126 in turn provides data communication services through the worldwide packet data communication network, now commonly referred to as the “Internet” 128. Local network 122 and Internet 128 both use electrical, electromagnetic, or optical signals that carry digital data streams. The signals through the various networks and the signals on network link 120 and through communication interface 118, which carry the digital data to and from computer system 100, are exemplary forms of carrier waves transporting the information.

[0101] Computer system 100 can send messages and receive data, including program code, through the network(s), network link 120, and communication interface 118. In the Internet example, a server 130 might transmit a requested code for an application program through Internet 128, ISP 126, local network 122 and communication interface 118. One such downloaded application may provide for the illumination optimization of the embodiment, for example. The received code may be executed by processor 104 as it is received, or stored in storage device 110, or other non-volatile storage for later execution. In this manner, computer system 100 may obtain application code in the form of a carrier wave.

[0102] While the concepts disclosed herein may be used for imaging on a substrate such as a silicon wafer, it shall be understood that the disclosed concepts may be used with any type of lithographic imaging systems, e.g., those used for imaging on substrates other than silicon wafers.

[0103] The terms “optimizing” and “optimization” as used herein refers to or means adjusting a patterning apparatus (e.g., a lithography apparatus), a patterning process, etc. such that results and/or processes have more desirable characteristics, such as higher accuracy of projection of a design pattern on a substrate, a larger process window, etc. Thus, the term “optimizing” and “optimization” as used herein refers to or means a process that identifies one or more values for one or more parameters that provide an improvement, e.g., a local optimum, in at least one relevant metric, compared to an initial set of one or more values for those one or more parameters. “Optimum” and other related terms should be construed accordingly. In an embodiment, optimization steps can be applied iteratively to provide further improvements in one or more metrics.

[0104] Embodiments of the present disclosure can be further described by the following clauses.

1. A method of generating a mask pattern, the method comprising:

obtaining a set of primitive elements of a target pattern;

encoding the set of primitive elements in an encoding space to generate a set of encoded primitive elements;

aggregating the set of encoded primitive elements in the encoding space to generate an encoded aggregation that represents the target pattern; and

generating a mask image based on the encoded aggregation, wherein the mask image is representative of a mask pattern for the target pattern.

2. The method of clause 1, wherein generating the mask image includes:

executing a mask prediction model that is configured to generate the mask image based on the encoded aggregation.

3. The method of clause 1, wherein aggregating the set of primitive elements includes:

executing a target representation model that is configured to generate the encoded aggregation based on the set of primitive elements, wherein the executing includes:

generating, by an encoder of the target representation model, the set of encoded primitive elements based on the set of primitive elements, and

generating, by an aggregator of the target representation model, the encoded aggregation based on the set of encoded primitive elements.

4. The method of clause 1 further comprising:

training a target representation model comprising an encoder and a decoder using a training set of primitive elements representing a training target pattern, wherein the encoder is configured to encode the training set of primitive elements to generate a training set of encoded primitive elements and the decoder is configured to decode the training set of encoded primitive elements to generate a decoded target pattern, wherein the target representation model is trained by comparing the decoded target pattern and the training target pattern.

5. The method of clause 4, wherein training the target representation model includes:

aggregating the training set of encoded primitive elements to generate a training encoded aggregation, and

decoding the training encoded aggregation to generate the decoded target pattern.

6. The method of clause 4, wherein training the target representation model includes:

generating multiple training target patterns using the training set of primitive elements.

7. The method of clause 6, wherein generating the multiple training target patterns includes:

deriving the training set of primitive elements from a training design layout generated based on a specified set of design rules, wherein the set of design rules are indicative of geometrical attributes associated with features of the training design layout.

8. The method of clause 4, wherein the target representation model and a mask prediction model that is configured to generate the mask image are trained simultaneously.

9. The method of clause 1, wherein obtaining the set of primitive elements includes:

decomposing a design layout to obtain a collection of primitive elements, wherein the decomposing includes identifying a feature, a portion of the feature, or a combination of features that occurs in multiple target patterns of the design layout as a primitive element of the collection of primitive elements, and wherein each target pattern of the target patterns is composed of one or more selected primitive elements of the collection of primitive elements in a specified arrangement, and

identifying the set of primitive elements from the collection of primitive elements for representing the target pattern.

10. The method of clause 9, wherein the set of primitive elements includes a parameterized primitive element, wherein the parameterized primitive element is associated with a parameter that is descriptive of an attribute of the feature and the parameter has different values for different instances of occurrence of the feature.

11. The method of clause 1, wherein the set of encoded primitive elements indicate locations of the set of primitive elements in the target pattern.

12. The method of clause 1, wherein each primitive element of the set of primitive elements is represented as an image.

13. The method of clause 12, wherein encoding the set of primitive elements includes: deriving a location of the primitive element in the target pattern from the image, and encoding the primitive element along with the location.

14. The method of clause 1, wherein each primitive element of the set of primitive elements is represented as a set of coordinates defining the corresponding primitive element.

15. The method of clause 14, wherein encoding the set of primitive elements includes: deriving a location of the primitive element in the target pattern from the set of coordinates, and encoding the primitive element along with the location.

16. The method of clause 1, wherein a primitive element of the set of primitive elements is represented as a set of attributes defining the primitive element.

17. The method of clause 16, wherein the primitive element is a circle, and wherein the set of attributes includes a radius of the circle and a coordinate of a center of the circle.

18. The method of clause 16, wherein the primitive element is a curvilinear feature, and wherein the set of attributes includes (a) coordinates of a set of points forming the curvilinear feature and (b) information regarding a tangent or directional derivative information for each point of the set of points.

19. The method of clause 1, wherein each primitive element of the set of primitive elements is represented as a set of elements defining the corresponding primitive element, wherein the set of elements includes at least one of vertices, edges, faces, or polygons.

20. The method of clause 1 further comprising:

training a mask prediction model, using (a) a training encoded aggregation corresponding to a training target pattern and (b) a training mask pattern corresponding to the training target pattern, to generate a predicted mask pattern, wherein the training encoded aggregation is generated by a target representation model, wherein the target representation model is trained, using a group of primitive elements corresponding to the training target pattern, to generate the training encoded aggregation.

21. The method of clause 20, wherein training the mask prediction model includes: generating a first target pattern and a first mask pattern corresponding to the first target pattern, wherein the first target pattern is generated by perturbing a group of primitive elements from the training target

pattern, and wherein the first mask pattern is generated by updating the training mask pattern in a portion corresponding to the group of primitive elements based on the perturbed group of primitive elements, and

training the mask prediction model using the first mask pattern and a first encoded aggregation that is generated for the first target pattern.

22. The method of clause 21, wherein training the mask prediction model includes:

generating multiple target patterns,

generating a set of aggregated encodings for the target patterns,

selecting a first target pattern from the multiple target patterns based on a distance metric between the set of aggregated encodings, and

training the mask prediction model using (a) a first encoded aggregation corresponding to the first target pattern and (b) a first mask pattern corresponding to the first target pattern.

23. The method of clause 1 further comprising:

determining a pattern coverage of a mask prediction model for predicting a first mask pattern for a first target pattern based on whether the first target pattern can be represented using primitive elements from a training set of primitive elements used in training a target representation model, wherein the target representation model is trained, using the training set of primitive elements of a training set of target patterns, to predict a training encoded aggregation for a training target pattern, and wherein the mask prediction model is trained to predict a training mask pattern based on the training encoded aggregation.

24. The method of clause 1 further comprising:

determining a prediction accuracy of a mask prediction model in predicting a first mask pattern for a first target pattern based on a pattern coverage of the mask prediction model, wherein the mask prediction model is trained, using a training set of encoded aggregation and a training set of mask patterns, to predict the mask pattern for the encoded aggregation of the target pattern, wherein the training set of encoded aggregation is generated by a target representation model that is trained, using a training set of primitive elements of a training set of target patterns, to predict the encoded aggregation for the target pattern.

25. The method of clause 24, wherein the pattern coverage is indicative of whether the first target pattern can be represented using the training set of primitive elements.

26. The method of clause 25, wherein the prediction accuracy is above a specified accuracy threshold based on the pattern coverage indicating that the first target pattern can be represented using the training set of primitive elements.

27. The method of clause 24, wherein the pattern coverage is characterized by using a difference between the first target pattern and a reconstructed first target pattern generated by a decoder of the target representation model using a first encoded aggregation of the first target pattern, wherein the first encoded aggregation is generated by the target representation model based on a first set of primitive elements of the first target pattern.

28. The method of clause 27, wherein the prediction accuracy is above a specified accuracy threshold based on the pattern coverage indicating that difference is below a specified threshold.

29. The method of clause 24, wherein the pattern coverage is characterized by a probability that a first encoded aggregation generated by the target representation model for the first target pattern belongs to a distribution of the training set of encoded aggregation in a feature vector space.

30. The method of clause 29, wherein the pattern coverage is characterized by a similarity metric of the first encoded aggregation, which is determined based on a distance between the first encoded aggregation and one or more encoded aggregation of the training set of encoded aggregation in a feature vector space.

31. The method of clause 30, wherein the prediction accuracy is above a specified accuracy threshold based on the pattern coverage indicating that the similarity metric is above a specified threshold.

32. An apparatus, the apparatus comprising:

a memory storing a set of instructions; and

a processor configured to execute the set of instructions to cause the apparatus to perform a method of any of the above clauses.

33. A non-transitory computer-readable medium having instructions recorded thereon, the instructions when executed by a computer implementing the method of any of the above clauses.

[0105] Aspects of the invention can be implemented in any convenient form. For example, an embodiment may be implemented by one or more appropriate computer programs which may be carried on an appropriate carrier medium which may be a tangible carrier medium (e.g., a disk) or an intangible carrier medium (e.g., a communications signal). Embodiments of the invention may be implemented using suitable apparatus which may specifically take the form of a programmable computer running a computer program arranged to implement a method as described herein. Thus, embodiments of the disclosure may be implemented in hardware, firmware, software, or any combination thereof.

Embodiments of the disclosure may also be implemented as instructions stored on a machine-readable medium, which may be read and executed by one or more processors. A machine-readable medium may include any mechanism for storing or transmitting information in a form readable by a machine (e.g., a computing device). For example, a machine-readable medium may include read only memory (ROM); random access memory (RAM); magnetic disk storage media; optical storage media; flash memory devices; electrical, optical, acoustical, or other forms of propagated signals (e.g., carrier waves, infrared signals, digital signals, etc.), and others. Further, firmware, software, routines, instructions may be described herein as performing certain actions. However, it should be appreciated that such descriptions are merely for convenience and that such actions in fact result from computing devices, processors, controllers, or other devices executing the firmware, software, routines, instructions, etc.

[0106] In block diagrams, illustrated components are depicted as discrete functional blocks, but embodiments are not limited to systems in which the functionality described herein is organized as illustrated. The functionality provided by each of the components may be provided by software or

hardware modules that are differently organized than is presently depicted, for example such software or hardware may be intermingled, conjoined, replicated, broken up, distributed (e.g., within a data center or geographically), or otherwise differently organized. The functionality described herein may be provided by one or more processors of one or more computers executing code stored on a tangible, non-transitory, machine-readable medium. In some cases, third party content delivery networks may host some or all of the information conveyed over networks, in which case, to the extent information (e.g., content) is said to be supplied or otherwise provided, the information may be provided by sending instructions to retrieve that information from a content delivery network.

[0107] Unless specifically stated otherwise, as apparent from the discussion, it is appreciated that throughout this specification discussions utilizing terms such as “processing,” “computing,” “calculating,” “determining” or the like refer to actions or processes of a specific apparatus, such as a special purpose computer or a similar special purpose electronic processing/computing device.

[0108] The reader should appreciate that the present application describes several inventions. Rather than separating those inventions into multiple isolated patent applications, these inventions have been grouped into a single document because their related subject matter lends itself to economies in the application process. But the distinct advantages and aspects of such inventions should not be conflated. In some cases, embodiments address all of the deficiencies noted herein, but it should be understood that the inventions are independently useful, and some embodiments address only a subset of such problems or offer other, unmentioned benefits that will be apparent to those of skill in the art reviewing the present disclosure. Due to cost constraints, some inventions disclosed herein may not be presently claimed and may be claimed in later filings, such as continuation applications or by amending the present claims. Similarly, due to space constraints, neither the Abstract nor the Summary sections of the present document should be taken as containing a comprehensive listing of all such inventions or all aspects of such inventions.

[0109] It should be understood that the description and the drawings are not intended to limit the present disclosure to the particular form disclosed, but to the contrary, the intention is to cover all modifications, equivalents, and alternatives falling within the spirit and scope of the inventions as defined by the appended claims.

[0110] Modifications and alternative embodiments of various aspects of the inventions will be apparent to those skilled in the art in view of this description. Accordingly, this description and the drawings are to be construed as illustrative only and are for the purpose of teaching those skilled in the art the general manner of carrying out the inventions. It is to be understood that the forms of the inventions shown and described herein are to be taken as examples of embodiments. Elements and materials may be substituted for those illustrated and described herein, parts and processes may be reversed or omitted, certain features may be utilized independently, and embodiments or features of embodiments may be combined, all as would be apparent to one skilled in the art after having the benefit of this description. Changes may be made in the elements described herein without departing from the spirit and scope of

the invention as described in the following claims. Headings used herein are for organizational purposes only and are not meant to be used to limit the scope of the description.

[0111] As used herein, unless specifically stated otherwise, the term “or” encompasses all possible combinations, except where infeasible. For example, if it is stated that a component includes A or B, then, unless specifically stated otherwise or infeasible, the component may include A, or B, or A and B. As a second example, if it is stated that a component includes A, B, or C, then, unless specifically stated otherwise or infeasible, the component may include A, or B, or C, or A and B, or A and C, or B and C, or A and B and C. Expressions such as “at least one of” do not necessarily modify an entirety of a following list and do not necessarily modify each member of the list, such that “at least one of A, B, and C” should be understood as including only one of A, only one of B, only one of C, or any combination of A, B, and C. The phrase “one of A and B” or “any one of A and B” shall be interpreted in the broadest sense to include one of A, or one of B.

[0112] The descriptions herein are intended to be illustrative, not limiting. Thus, it will be apparent to one skilled in the art that modifications may be made as described without departing from the scope of the claims set out below.

CLAIMS

1. A method of generating a mask pattern, the method comprising:

obtaining a set of primitive elements of a target pattern;

5 encoding the set of primitive elements in an encoding space to generate a set of encoded primitive elements;

aggregating the set of encoded primitive elements in the encoding space to generate an encoded aggregation that represents the target pattern; and

10 generating a mask image based on the encoded aggregation, wherein the mask image is representative of a mask pattern for the target pattern.

2. The method of claim 1, wherein generating the mask image includes:

executing a mask prediction model that is configured to generate the mask image based on the encoded aggregation.

15

3. The method of claim 1, wherein aggregating the set of primitive elements includes:

executing a target representation model that is configured to generate the encoded aggregation based on the set of primitive elements, wherein the executing includes:

generating, by an encoder of the target representation model, the set of encoded primitive elements

20 based on the set of primitive elements, and

generating, by an aggregator of the target representation model, the encoded aggregation based on the set of encoded primitive elements.

4. The method of claim 1 further comprising:

25 training a target representation model comprising an encoder and a decoder using a training set of primitive elements representing a training target pattern, wherein the encoder is configured to encode the training set of primitive elements to generate a training set of encoded primitive elements and the decoder is configured to decode the training set of encoded primitive elements to generate a decoded target pattern, wherein the target representation model is trained by comparing the decoded target pattern and the training target pattern,

30

wherein training the target representation model includes:

aggregating the training set of encoded primitive elements to generate a training encoded aggregation, and

decoding the training encoded aggregation to generate the decoded target pattern.

35

5. The method of claim 3, wherein training the target representation model includes:

generating multiple training target patterns using the training set of primitive elements,

wherein generating the multiple training target patterns includes:

deriving the training set of primitive elements from a training design layout generated based on a specified set of design rules, wherein the set of design rules are indicative of geometrical attributes associated with features of the training design layout.

5

6. The method of claim 1, wherein obtaining the set of primitive elements includes:

decomposing a design layout to obtain a collection of primitive elements, wherein the decomposing includes identifying a feature, a portion of the feature, or a combination of features that occurs in multiple target patterns of the design layout as a primitive element of the collection of primitive

10 elements, and wherein each target pattern of the target patterns is composed of one or more selected primitive elements of the collection of primitive elements in a specified arrangement, and identifying the set of primitive elements from the collection of primitive elements for representing the target pattern.

15 7. The method of claim 6, wherein the set of primitive elements includes a parameterized primitive element, wherein the parameterized primitive element is associated with a parameter that is descriptive of an attribute of the feature and the parameter has different values for different instances of occurrence of the feature.

20 8. The method of claim 1, wherein the set of encoded primitive elements indicate locations of the set of primitive elements in the target pattern.

9. The method of claim 1, wherein each primitive element of the set of primitive elements is represented as an image,

25 wherein encoding the set of primitive elements includes:

deriving a location of the primitive element in the target pattern from the image, and encoding the primitive element along with the location.

10. The method of claim 1, wherein each primitive element of the set of primitive elements is represented as a set of coordinates defining the corresponding primitive element,

30 wherein encoding the set of primitive elements includes:

deriving a location of the primitive element in the target pattern from the set of coordinates, and encoding the primitive element along with the location.

35 11. The method of claim 1, wherein a primitive element of the set of primitive elements is represented as a set of attributes defining the primitive element,

wherein the primitive element is a curvilinear feature, and wherein the set of attributes includes (a) coordinates of a set of points forming the curvilinear feature and (b) information regarding a tangent or directional derivative information for each point of the set of points.

5 12. The method of claim 1, wherein each primitive element of the set of primitive elements is represented as a set of elements defining the corresponding primitive element, wherein the set of elements includes at least one of vertices, edges, faces, or polygons.

13. The method of claim 1 further comprising:
10 training a mask prediction model, using (a) a training encoded aggregation corresponding to a training target pattern and (b) a training mask pattern corresponding to the training target pattern, to generate a predicted mask pattern, wherein the training encoded aggregation is generated by a target representation model, wherein the target representation model is trained, using a group of primitive elements corresponding to the training target pattern, to generate the training encoded aggregation.

15 14. The method of claim 13, wherein training the mask prediction model includes:
generating a first target pattern and a first mask pattern corresponding to the first target pattern, wherein the first target pattern is generated by perturbing a group of primitive elements from the training target pattern, and wherein the first mask pattern is generated by updating the training mask
20 pattern in a portion corresponding to the group of primitive elements based on the perturbed group of primitive elements, and
training the mask prediction model using the first mask pattern and a first encoded aggregation that is generated for the first target pattern.

25 15. The method of claim 13, wherein training the mask prediction model includes:
generating multiple target patterns,
generating a set of aggregated encodings for the target patterns,
selecting a first target pattern from the multiple target patterns based on a distance metric between the set of aggregated encodings, and
30 training the mask prediction model using (a) a first encoded aggregation corresponding to the first target pattern and (b) a first mask pattern corresponding to the first target pattern.

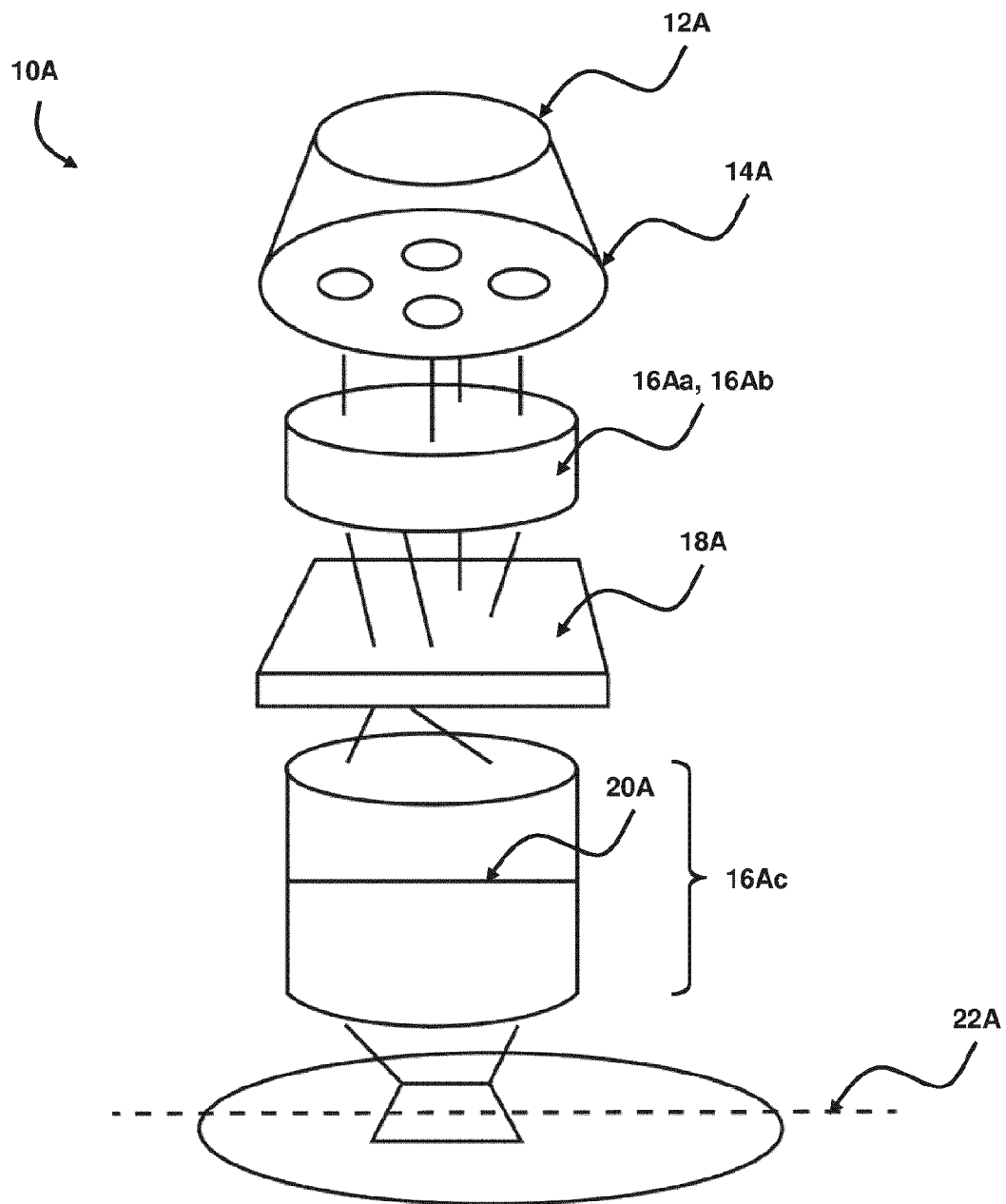


FIG. 1

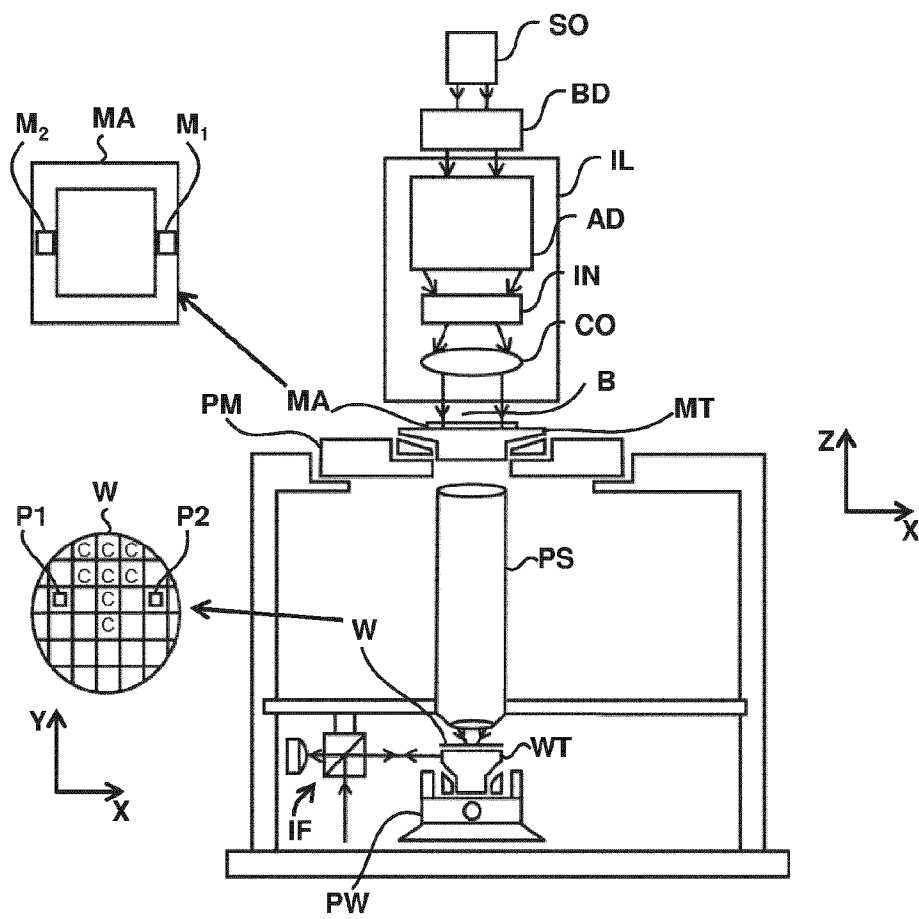


FIG. 2

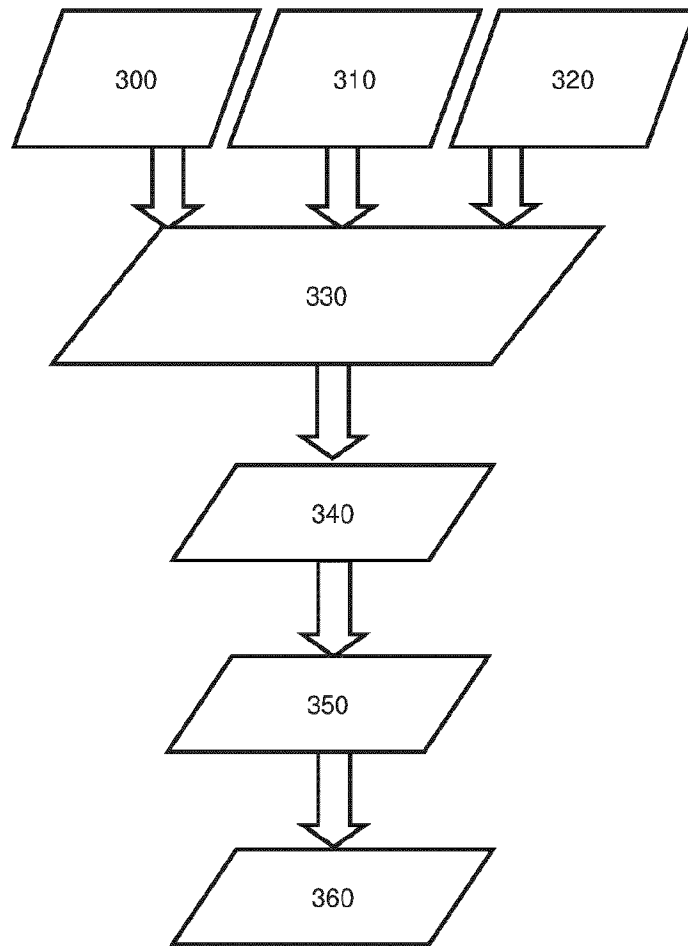


FIG. 3

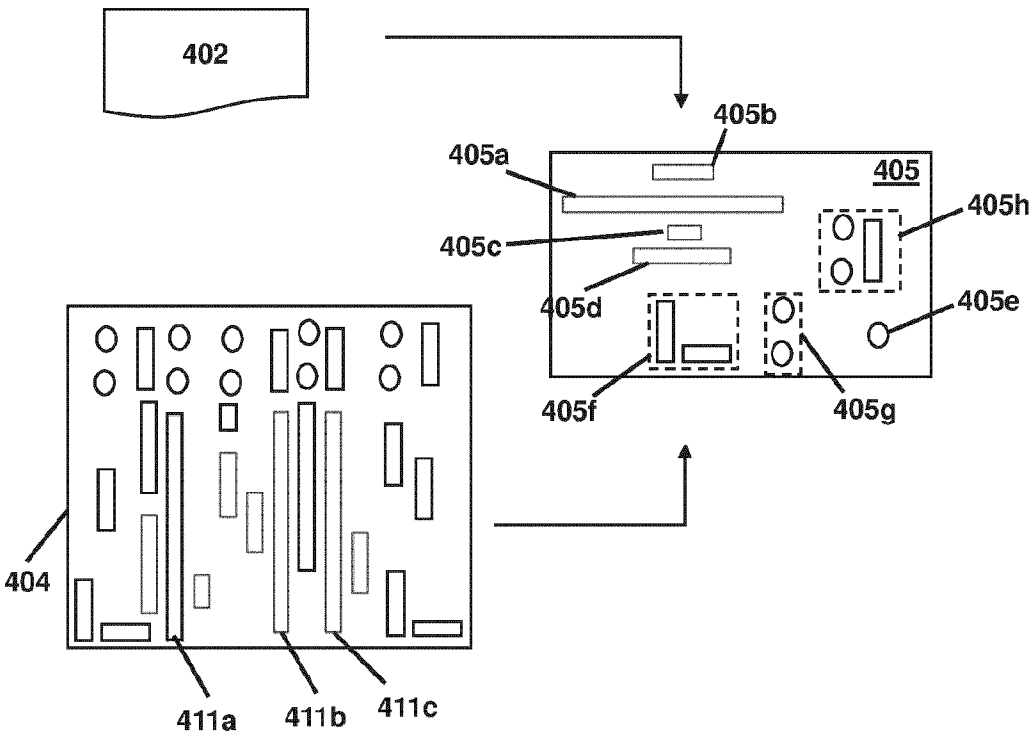


FIG. 4A

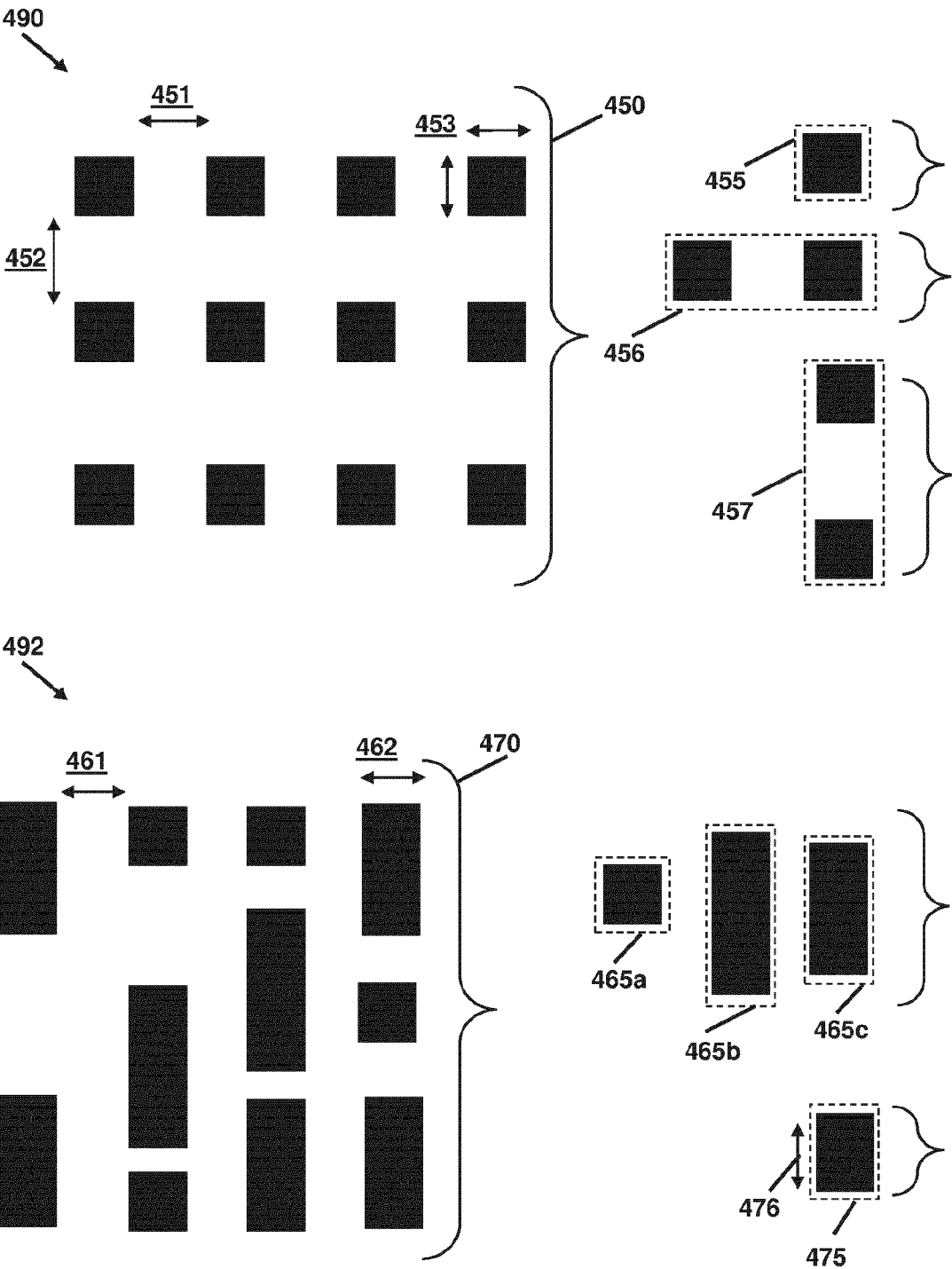


FIG. 4B

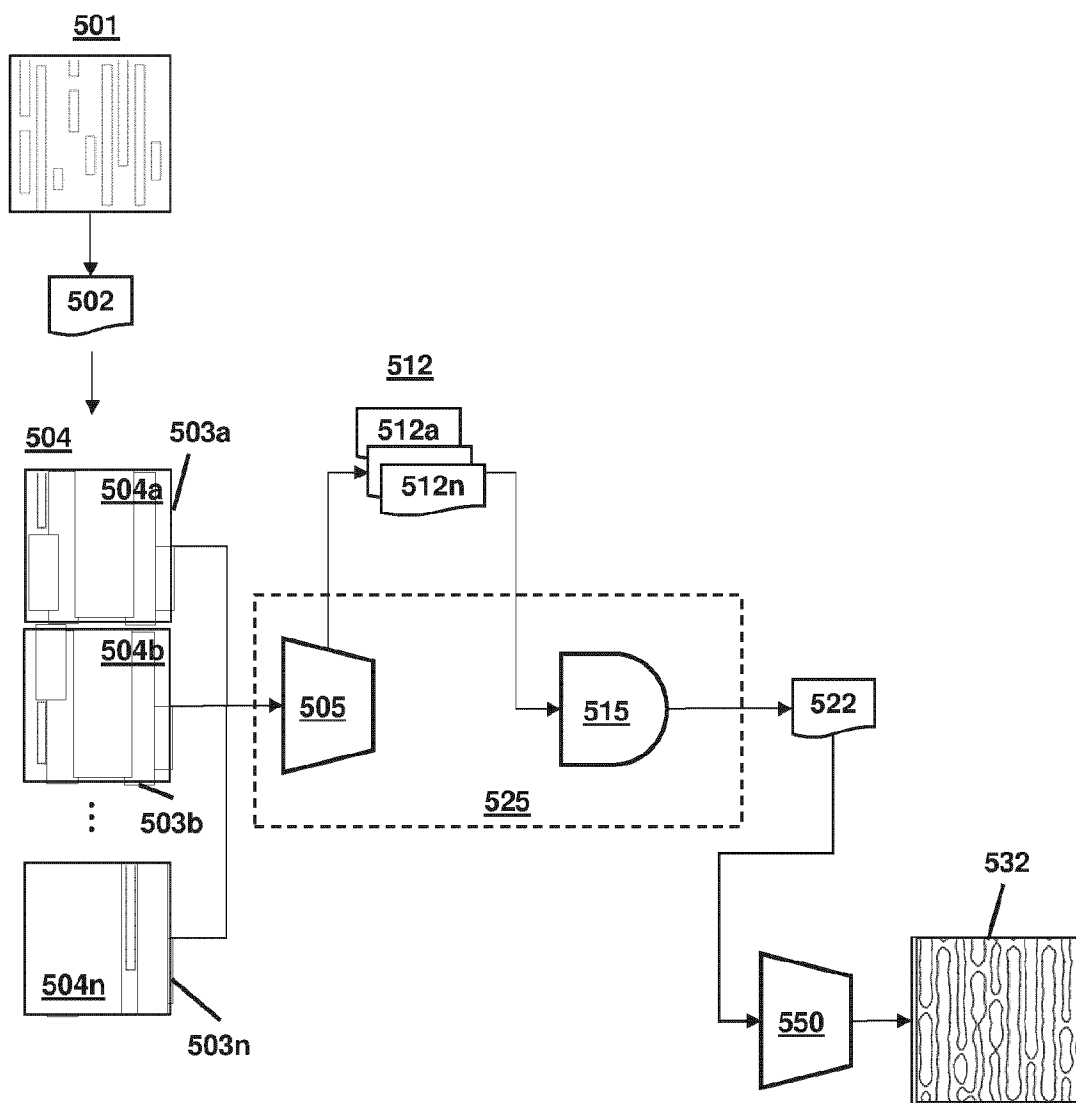


FIG. 5

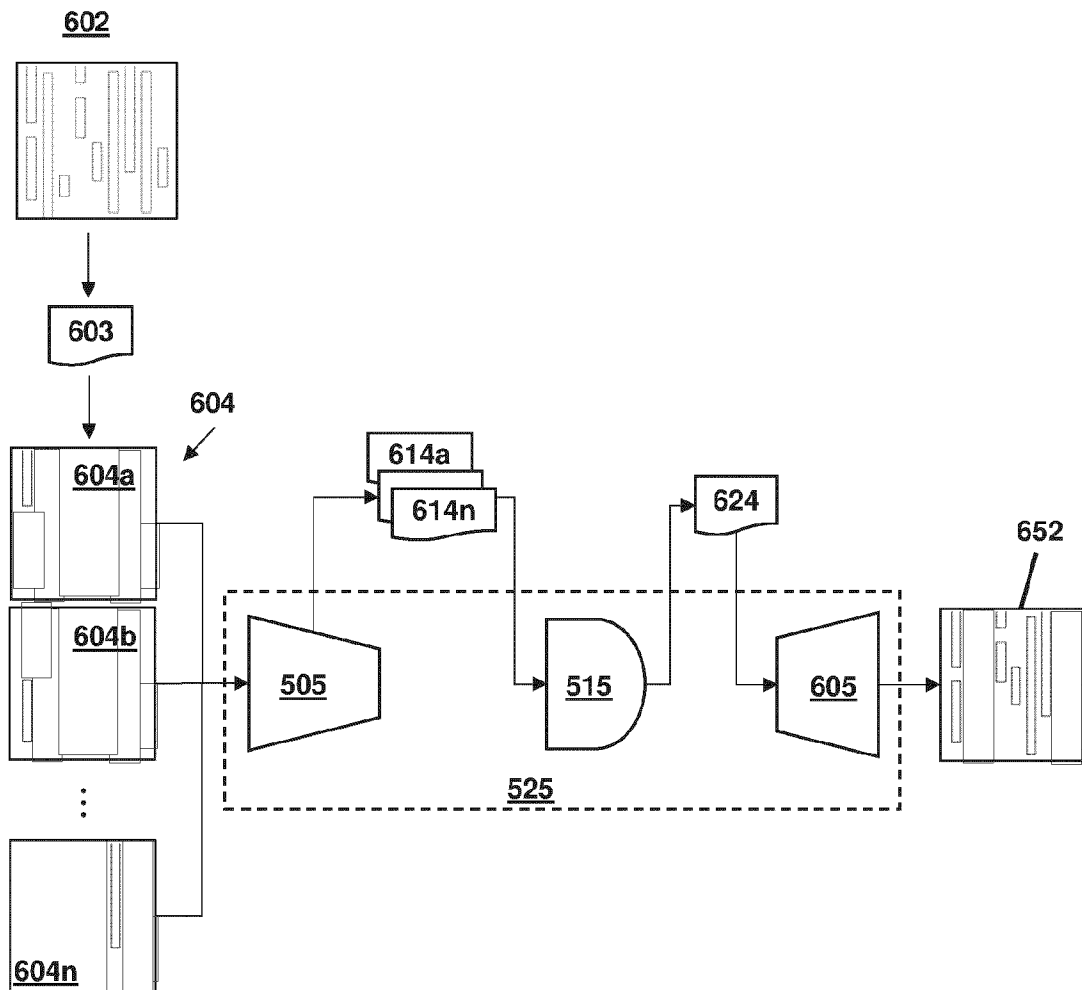


FIG. 6

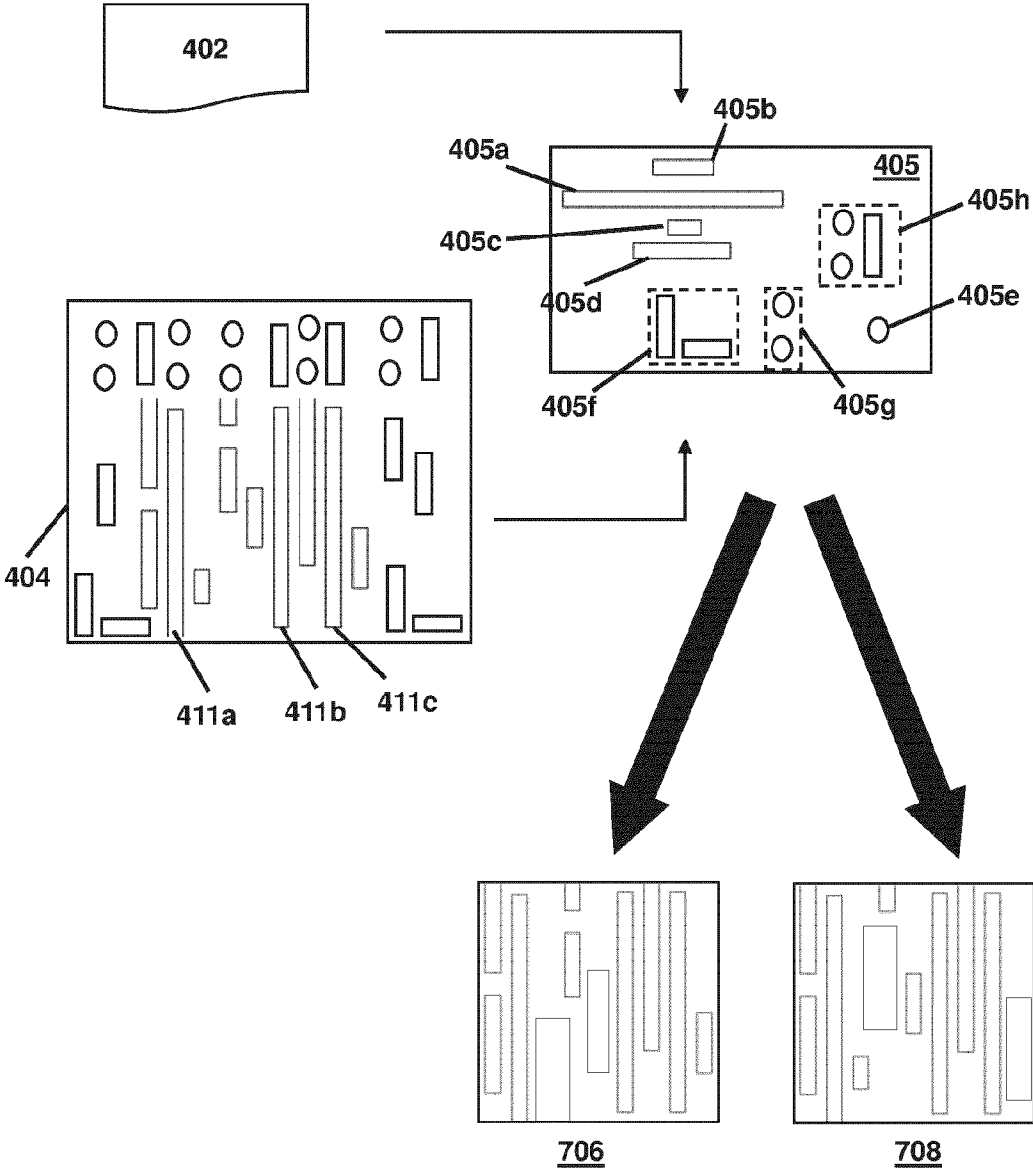


FIG. 7

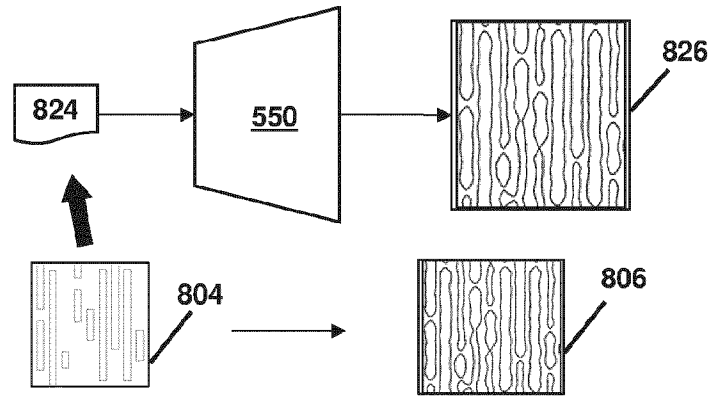


FIG. 8

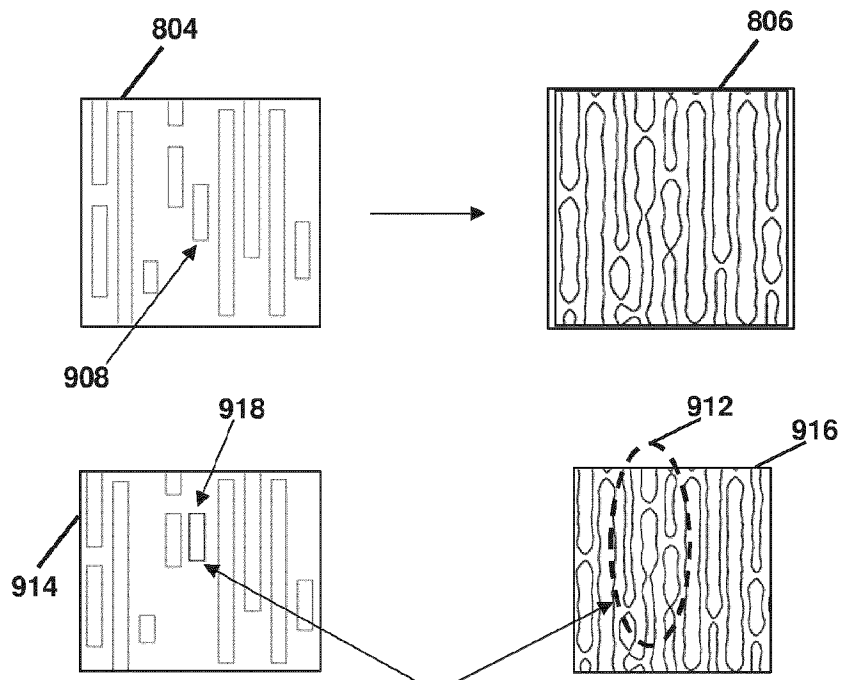


FIG. 9

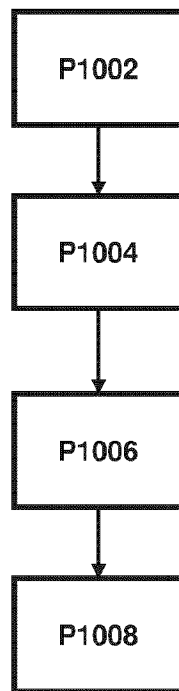


FIG. 10

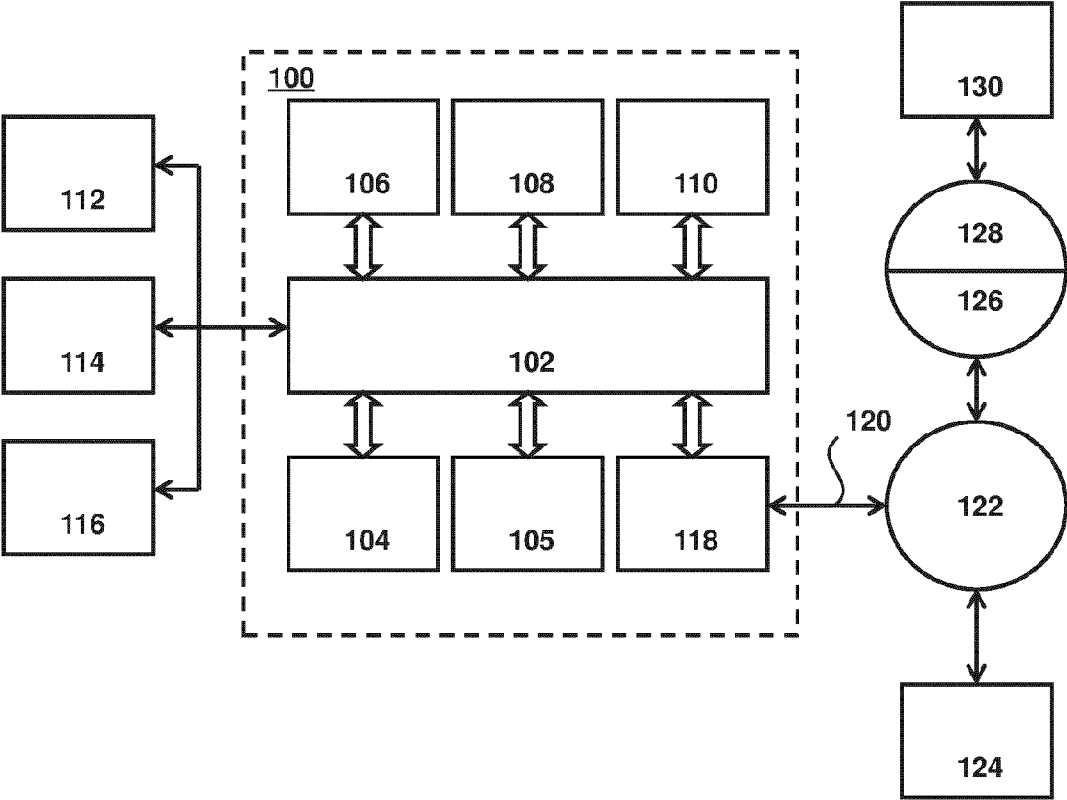


FIG. 11

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2024/082675

A. CLASSIFICATION OF SUBJECT MATTER INV. G03F1/36 G03F7/00 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) G03F Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 2022/263312 A1 (ASML NETHERLANDS BV [NL]) 22 December 2022 (2022-12-22) paragraphs [0006] - [0010], [0061] - [0068], [0093] - [0105]; figure 5 -----	1 - 15
A	US 2022/128899 A1 (FUJIMURA AKIRA [US] ET AL) 28 April 2022 (2022-04-28) paragraphs [0008], [0026], [0072] - [0078]; figures 9-11 -----	1 - 15
A	WO 2023/180020 A1 (ASML NETHERLANDS BV [NL]) 28 September 2023 (2023-09-28) paragraphs [0019], [0070] - [0078]; figures 3-5 -----	1 - 15
☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 21 March 2025		Date of mailing of the international search report 02/04/2025
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Roesch, Guillaume

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2024/082675

Patent document cited in search report			Publication date		Patent family member(s)		Publication date	
WO 2022263312	A1	22-12-2022	CN	117480449	A		30-01-2024	
			KR	20240023521	A		22-02-2024	
			TW	202318097	A		01-05-2023	
			TW	202419964	A		16-05-2024	
			US	2024256976	A1		01-08-2024	
			WO	2022263312	A1		22-12-2022	

US 2022128899	A1	28-04-2022	EP	4205176	A1		05-07-2023	
			JP	2023547622	A		13-11-2023	
			KR	20230091155	A		22-06-2023	
			TW	202223547	A		16-06-2022	
			TW	202343159	A		01-11-2023	
			US	2022128899	A1		28-04-2022	
WO 2023180020	A1	28-09-2023	WO	2022086825	A1		28-04-2022	

			CN	118829944	A		22-10-2024	
			EP	4497035	A1		29-01-2025	
			KR	20240163644	A		19-11-2024	
			TW	202407459	A		16-02-2024	
WO	2023180020	A1		28-09-2023				
