(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
06 February 2020 (06.02.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/028302 A1

(51) International Patent Classification:
E21B 7/04 (2006.01) *G06Q 50/02* (2012.01)

(21) International Application Number:
PCT/US2019/044042

(22) International Filing Date:
30 July 2019 (30.07.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/712,518 31 July 2018 (31.07.2018) US
18194788.8 14 September 2018 (14.09.2018) EP

(71) Applicant *(for US only)*: **SHELL OIL COMPANY** [US/US]; P.O. Box 576, Houston, Texas 77001-0576 (US).

(71) Applicant *(for all designated States except US)*: **SHELL INTERNATIONALE RESEARCH MAATSCHAPPIJ B.V.** [NL/NL]; Carel van Bylandtlaan 30, 2596 The Hague (NL).

(72) Inventors: **PANCHAL, Neilkunal**; 3333 Hwy 6, Houston, Texas 77082 (US). **SULTAN, Sami Mohammed Khair**; 20102 Echo Falls Dr, Katy, Texas 77480 (US). **VILA, Jeremy Paul**; 2222 Smith St. Unit 301, Houston, Texas 77002 (US). **JAIN, Minith Bharat**; 457, 3rd Flr, 304, Satyam Bldg, Nazrana Comp, Kasar Ali, Bhiwandi, Thane, Maharashtra 421302 (IN). **THANOON, David**; 2304 McClendon St, Houston, Texas 77030 (US). **UZCATEGUI DIAZ, Misael Jacobo**; 14220 Park Row Apt 311, Houston, Texas 77084 (US). **CHATTERJEE, Arnab**; 1180A Eddy St, San Francisco, California 94109 (US).

(74) Agent: **VANDENHOFF, Deborah**; SHELL OIL COMPANY, P.O. Box 576, Houston, Texas 77001-0576 (US).

(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, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, 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, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, 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,

(54) Title: PROCESS FOR REAL TIME GEOLOGICAL LOCALIZATION WITH REINFORCEMENT LEARNING

(57) Abstract: A method of geosteering in a wellbore construction process uses an earth model that defines boundaries between formation layers and petrophysical properties of the formation layers in a subterranean formation. Sensor measurements related to the wellbore construction process are inputted to the earth model. An estimate is obtained for a relative geometrical and geological placement of the well path with respect to a geological objective using a trained reinforcement learning agent. An output action based on the sensor measurement for influencing a future profile of the well path with respect to the estimate.

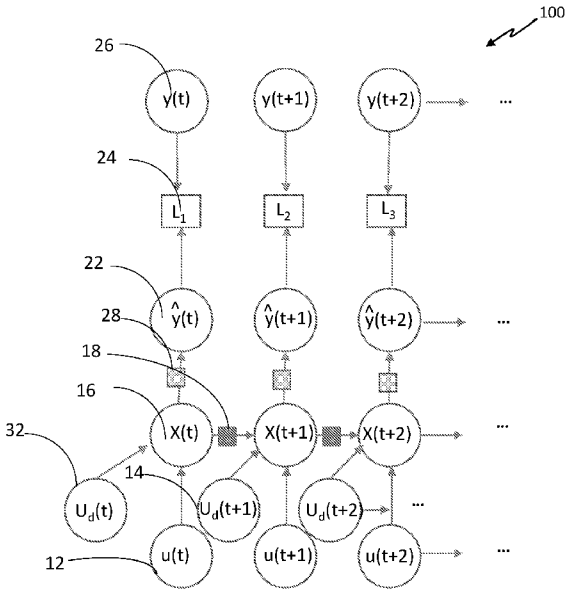


FIG. 1

WO 2020/028302 A1

[Continued on next page]

EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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).

Published:

— *with international search report (Art. 21(3))*

— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*

PROCESS FOR REAL TIME GEOLOGICAL LOCALIZATION WITH REINFORCEMENT LEARNING

FIELD OF THE INVENTION

[0001] The present invention relates to the field of geosteering and, in particular, to a process for real time geological localization with reinforcement learning for automating parts of a geological steering workflow.

BACKGROUND OF THE INVENTION

[0002] In a well construction process, rock destruction is guided by a drilling assembly. The drilling assembly includes sensors and actuators for biasing the trajectory and determining the heading in addition to properties of the surrounding borehole media. The intentional guiding of a trajectory to remain within the same rock or fluid and/or along a fluid boundary, such as an oil/water contact or an oil/gas contact, is known as geosteering.

[0003] The objective in drilling wells is to maximize the drainage of fluid in a hydrocarbon reservoir. Multiple wells placed in a reservoir are either water injector wells or producer wells. The objective is maximizing the contact of the wellbore trajectory with geological formations that: are more permeable, drill faster, contain less viscous fluid, and contain fluid of higher economical value. Furthermore, drilling more tortuous wells, slower, and out of zone add to the costs of the well.

[0004] Geosteering is drilling a horizontal wellbore that ideally is located within or near preferred rock layers. As interpretive analysis is performed while or after drilling, geosteering determines and communicates a wellbore's stratigraphic depth location in part by estimating local geometric bedding structure. Modern geosteering normally incorporates more dimensions of information, including insight from downhole data and quantitative correlation methods. Ultimately, geosteering provides explicit approximation of the location of nearby geologic beds in relationship to a wellbore and coordinate system.

[0005] Geosteering relies on mapping data acquired in the structural domain along the horizontal wellbore and into the stratigraphic depth domain. Relative Stratigraphic Depth (RSD) means that the depth in question is oriented in the stratigraphic depth direction and is relative to a geologic marker. Such a marker is typically chosen from type log data to be the top of the pay zone/target layer. The actual drilling target or “sweet spot” is located at an onset stratigraphic distance from the top of the pay zone/target layer.

[0006] In an article by H. Winkler ("Geosteering by Exact Inference on a Bayesian Network" Geophysics 82:5:D279-D291; Sept-Oct 2017), machine learning is used to solve a Bayesian network. For a sequence of log and directional survey measurements, and a pilot well log representing a geologic column, a most likely well path and geologic structure is determined.

[0007] There remains a need for autonomous geosteering processes with improved accuracy.

SUMMARY OF THE INVENTION

[0008] According to one aspect of the present invention, there is provided a method of geosteering in a wellbore construction process, the method comprising the steps of: providing an earth model defining boundaries between formation layers and petrophysical properties of the formation layers in a subterranean formation comprising data selected from the group consisting of seismic data, data from an offset well and combinations thereof; comparing sensor measurements related to the wellbore construction process to the earth model; obtaining an estimate from the earth model for a relative geometrical and geological placement of the well path with respect to a geological objective using a trained reinforcement learning agent; and determining an output action based on the sensor measurement for influencing a future profile of the well path with respect to the estimate.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The method of the present invention will be better understood by referring to the following detailed description of preferred embodiments and the drawings referenced therein, in which:

[00010] Fig. 1 is a flow diagram illustrating one embodiment of the method of the present invention;

[00011] Fig. 2 illustrates one embodiment of a work flow of the method of the present invention;

[00012] Fig. 3 illustrates another embodiment of a work flow of the method of the present invention;

[00013] Fig. 4 is a graphical representation of the results of a first test of a simulation environment produced according to the method of the present invention;

[00014] Fig. 5 is a graphical representation of the results of a second test of a simulation environment produced according to the method of the present invention;

[00015] Fig. 6 is a graphical representation of the results of a third test of a simulation environment produced according to the method of the present invention; and

[00016] Fig. 7 is a graphical representation of the results of a fourth test of a simulation environment produced according to the method of the present invention

DETAILED DESCRIPTION OF THE INVENTION

[00017] The present invention provides a method for geosteering in a wellbore construction process. A wellbore construction process can be a wellbore drilling process. The method is advantageously conducted while drilling. The method uses a trained reinforcement learning agent. The method is a computer-implemented method.

[00018] In accordance with the present invention, an earth model is provided. The earth model defines boundaries between formation layers and petrophysical properties of the formation layers of a subterranean formation. The earth model is produced from data relating to a subterranean formation, the data selected from the group consisting of seismic data, data from an offset well and combinations thereof. Preferably, the earth model is a 3D model.

[00019] The earth model may be a static or dynamic model. Preferably, the earth model is a dynamic model that changes dynamically during the drilling process.

[00020] Sensor measurements are inputted to the earth model. The sensor measurements are obtained during the wellbore construction process. Accordingly, real-time sensor measurements are made while drilling. In a real-time drilling process, sensors are chosen based on the geological objectives. if the target reservoir and the surrounding medium can be distinguished by a particular measurement, then this measurement will be chosen. Since there is a limit of the telemetry rate, the sample frequency would also be budgeted. Preferably, the sensor measurements are provided as a streaming sequence. The sensors may be LWD sensors, MWD sensors, image logs, 2D seismic data, 3D seismic data and combinations thereof.

[00021] The LWD sensor may be selected from the group consisting of gamma-ray detectors, neutron density sensors, porosity sensors, sonic compressional slowness sensors, resistivity sensors, nuclear magnetic resonance, and combinations thereof.

[00022] The MWD sensor is selected from the group consisting of sensors for measuring mechanical properties, inclination, azimuth, roll angles, and combinations thereof.

[00023] The earth model simulates the earth and then a sensor measurement from the earth. The simulated sensor measurement is then compared to an actual sensor measurement made while drilling.

[00024] A well path is selected to reach a geological objective, such as a geological feature, such as fault, a nearby offset well, a fluid boundary and the like. Examples of fluid boundaries may be oil/water contacts, oil/gas contacts, oil/tar contacts, and the like. An estimate for the relative geometrical and geological placement of a well path to reach the geological objective is obtained using a trained reinforcement learning agent. An output action based on the sensor measurement for influencing a future profile of the well path is determined with respect to the estimate.

[00025] The trained reinforcement learning agent is preferably a trained Bayesian reinforcement learning (BRL) agent or a trained Monte Carlo Trajectory Sampling (MCTS) reinforcement learning agent.

[00026] Preferably, a component of the trained BRL agent is a Markov Decision Process (MDP). The data used for training may be historical or synthetic data.

[00027] The trained MCTS reinforcement learning agent is defined with respect to a distribution over RSD transitions where the distribution is determined from a Monte Carlo Tree Search.

[00028] In a preferred embodiment, the output action of the reinforcement learning agent is determined by maximizing the placement of the well path with respect to a geological datum. An objective is maximizing the contact of the wellbore trajectory with geological formations that: are more permeable, drill faster, contain less viscous fluid, and contain fluid of higher economical value. The geological datum can be, for example, without limitation, a rock formation boundary, a geological feature, an offset well, an oil/water contact, an oil/gas contact, an oil/tar contact and combinations thereof.

[00029] The steering of the wellbore trajectories is achieved through a number of different actuation mechanisms, including, for example, rotary steerable systems (RSS) or positive displacement motors. The former contains downhole actuation, power generation feedback control and sensors, to guide the bit by either steering an intentional bend in systems known as point-the-bit or by applying a sideforce in a push-the-bit system. PDM motors contain a

fluid actuated Moyno motor that converts hydraulic power to rotational mechanical power for rotating a bit. the motor contains a bend such that the axis of rotation of the bit is offset from the centerline of the drilling assembly. Curved boreholes are achieved through circulating fluid through the motor and keeping the drill-string stationary. Curved boreholes are achieved through rotating the drill string whilst circulating such that the bend cycle averages to obtain a straight borehole.

[00030] The output action can be curvature, roll angle, set points for inclination, set points for azimuth, Euler angle, rotation matrix quaternions, angle axis, position vector, position Cartesian, polar, and combinations thereof.

[00031] In a preferred embodiment, the estimate for a relative geometrical and geological placement of the well path is determined by providing to the trained reinforcement learning agent a state space representation for a given depth for a position and a direction of the well path and the geological datum, having a discretized representation of the output action as a set of plausible geological datum changes; a state transition function for determining a transition between the state space representation at depth t and depth t+1 conditional upon the output action; an observational model for modeling the sensor measurements to the earth model; a reward function; a discount rate applied to the reward function for determining a discounted reward function; and a value function representing a past sum of discounted rewards for the transition of depth running forward in time.

[00032] Referring now to Fig. 1, a flow diagram of the method 100 of one embodiment of the present invention is illustrated, where the trained reinforcement learning agent is a BRL agent. As illustrated, u(t) is the control vector 12 for influencing the drilling process. An example of the control vector 12 may be u_{tf} toolface angle, an inclination set point, and the like. Edge 14 representing the mapping from control to state. x(t), denoted by reference numeral 16, is the state vector such as position and heading. A state transition model 18 is represented by $P(x(t+1)|x(t))$, where $x(t+1) = f(x(t), u(t))$ is a dynamical system response. The predicted output 22 simulates a sensor measurement. The method 100 includes a loss function or error function 24, sequential data 26, an observation model 28 and a formation interpretation decision variable 32.

[00033] Fig. 2 illustrates one embodiment of a work flow for drilling, where the trained reinforcement learning agent is a BRL agent.

[00034] Preferably, an optimal output action for a most probable well path is solved with respect to the value function to minimize or maximize the expected sum of the reward function at a given depth. An optimum value function is determined by iterating on a maximum or minimum of the expected sum of the reward function at depth t with the value of the state space at depth t-1 with respect to state transition function, selecting the highest value state with respect to a constraint, and propagating forward in depth the output actions to determine an optimum formation interpretation.

[00035] Preferably, the state space is continuous.

[00036] In a preferred embodiment, the state transition function is pretrained on historical wells and or synthetic data. The function may be trained on a neural network and/or a probabilistic graphical model. The probabilistic graphical model may be a Dirichlet-multinomial exponential family conjugate prior representation where the hyper parameters are trained by counting state visits.

[00037] Preferably, the discounted sum of rewards is based on discretized depth intervals in an arc length of the well path. The reward function is selected from the group consisting of a sequence similarity measure, a mean squared error reward function, a Huber loss reward function, a non-convex reward function and combination thereof.

[00038] Preferably, the observation model is a look-up from a type log or the earth model.

[00039] In accordance with the present invention, a propagating borehole accumulates arclength $s \in R$. It is assumed that the geosteering problem is reduced to a 2D problem because of the horizontal rock layers. In this 2D section the position of the bit is defined as $x = (x_{tvd}, x_{xsec}) \in R^2$, where x_{tvd} is the true vertical depth defined to be positive vertically and x_{xsec} is the vertical departure in the 2D cross section.

[00040] The formations are assumed to be parallel and unfaulted. The top of a given reference formation of interest is given as $x_f(x_{xsec})$. When geosteering with respect to a reference formation, the relative stratigraphic distance defined with respect to this formation is defined to be $x_{rsd} = x_{tvd}(x_{xsec}) - x_f(x_{xsec})$. The geosteering problem of this paper refers to a single measurement with respect to an offset reference well. The measurement of the reference well is denoted $\gamma_t(x_{rsd})$ representing gamma ray counts with respect to relative stratigraphic depth x_{rsd}. The propagating drilling assembly measures the surrounding medium and returns a measurement $\gamma_w(s)$. The objective of the geosteering optimization problem of

this paper is to determine the relative stratigraphic position of the wellbore $x_{rsd}(x_{xsec})$, from the observations $\gamma_w(s)$ with respect to the reference well $\gamma_t(x_{rsd})$.

[00041] A Markov decision process (X,U,P,R,γ) is a 5-tuple where S is a set of states, A is a finite set of actions, $P(x(t+1) \mid x(t),u(t))$ is the probability that state x(t) at time t and control u(t) will lead to state x(t + 1) at time t + 1, R(x(t + 1), x(t), u(t)) is the reward from transitioning from state x(t) at time t to state x(t+1) at time t+1 due to control action u(t) and $\gamma \epsilon [0,1]$ is the discount rate.

[00042] The goal of an MDP problem is to find a policy π(x(t)) that minimizes a value function V(x(t)) where

$$V(x(t)) = \sum_{t=0}^{t=\infty} r(x((t), x((t+1)u(t))$$

and to choose an action $u(t) \epsilon \pi(x(t))$ that maximizes the value function V(x(t))

$$u(t) = \pi(x(t)) = \arg\max \pi \sum_{t=0}^{t=\infty} r(x((t), x((t+1)u(t))$$

[00043] In a geosteering problem, a Markov decision process has a state space defined as $X = \{x_{rsd}(t) \epsilon x_{rsd0}, x_{rsd1}, \ldots, x_{rsdn}\}$ where $n \epsilon N$ to be a finitely spaced discrete set of positions representing stratigraphic distances relative to a formation boundary.

$$x_{rsd}(t=1) = x_{rsd}(t) + \beta\left(u_{fdip}(t) - u_{inc}(t)\right) + \eta$$

where the noise is normally distributed

$$\eta \sim N(\mu, \sigma^2)$$

with mean μ and variance σ^2. The formation dip angle is denoted by $u_{dip}(t) \epsilon (0,\pi)$ and the inclination angle is $u_{inc}(t) \epsilon (0,\pi)$. For the geosteering problems, the inclination angle is known albeit noisy and the goal of a geosteerer is to determine the sequence of dip angles to determine the relative stratigraphic position trajectory in the real-tie process.

$$x_{rsd}(t+d) = x_{rsd}(t) + u_{dip}(t) + \eta$$

where $u_{dip}(t) = \beta\left(u_{fdip}(t) - u_{inc}(t)\right) \in U$.

[00044] For learning the Markov decision process, assume a state transition function which represents the system dynamics given by:

$$P(x(t+d) \mid x(t), u(t)) = \eta$$
$$x(t+d) = x(t) + u_d + \eta$$

[00045] This is a simple linear approximation of the real dynamics. Once learned this model can be used to optimize the formation interpretation algorithm in the dynamic programming step.

$$V\big(x(t)\big) = max_a E\{r\big(x(t), u(t)\big) + \backslash gammaV(x(t+1))\}$$

[00046] Given a discretization of x(t), u(t) and x(t+d) into a user defined finite interval, where u(t) is the decision variable, then the dynamics $P\big(x(t+d)\big|x(t), u(t)\big)$ can be thought of as a 3D array. The first step to learning this model is to process historical data into a buffer such that for a given *t* and *d* each state, control and next state is lined up. Make sure that d is appropriately chosen for if it is too small than the state transitions will not be captured, if too large than the linear approximation to the dynamics no longer become valid.

[00047] A naïve way of learning this model is to update the count in the table corresponding to the triple for each transition encountered in the data and then to normalize the row across the x(t+d) row. A better way to learn this model using Bayesian methods would be to assign a prior Dirichlet distribution $D(\alpha_i)$ for each state and control x(t) and u(t) to represent the distribution over next states *x(t+d)*.

[00048] This provides a prior distribution of the entire system dynamics by productizing over all controls and states such that

$$P\big(x(t+d)\big|x(t), u(t)\big) = \prod_{i,j} D(\alpha_i) D(\alpha_j)$$

[00049] Where D is the Dirichlet distribution. Here the Dirichlet distribution is conjugate to the multinomial distribution, and hence a multinomial distribution can be fitted to consecutive data points in the buffer to update the counts of the α vector.

[00050] This method can be extended to other exponential family distributions explaining the state transition function in similar ways. Furthermore, the state transition can be compounded to be a mixture of multinomial distributions.

[00051] The method can also extend for nonexponential family distributions where either sampling methods can be used. Alternatively, a neural network function approximator can be used. In one embodiment, a forward pass from a state x(t) and control u(t) vector once concatenated passes through a sequence of neurons represented by affine transformations followed by non-linear activation functions, such as “relu”, “selu”, “tanh” etc., until a final fully connected layer where an output of *x(t+d)* exists. This is paired with a loss function, often a mean squatted error, although other functions such as Huber norm, L1 norm can be

used for regression or softmax with cross entropy for categorical distributions as is the case for this embodiment.

[00052] By sampling a batch from the buffer and performing a forward pass, back propagation of the derivatives can use used to optimize the weights and biases of the affine transformation components in the neural network model. Once trained with sufficient data. The trained model can be used in the dynamic programming or optimization steps.

[00053] The buffer can be created from historical data or from simulated data. To validate the model, a sample of data where the resultant state transition x(t+d) in known is used to compare the predicted state transition x(t+d) from the real. This also serves to determine if the training is over fitting and under fitting and if regularization techniques need to be employed. Preferably, a reward function is selected to maximize the similarity between a sensor measurement sequence $\gamma_w(t_0 : t_0 + d) = (\gamma_w(t_0), \gamma_w(t_0 + 1), \ldots, \gamma_w(t_0 + d))$ over a fixed length interval $d \in N$ and a sequence generated by the model $x_{rsd}(t + d) = x_{rsd}(t) + u_{dip}(t) + \eta,\ \gamma_t(x_{rsd}(t_0 : t_0 + d)) = (\gamma_t(x_{rsd}(t_0)), \gamma_t(x_{rsd}(t_0 + 1)), \ldots, \gamma_t(x_{rsd}(t_0 + d)))$

$$r\big(x(t), x(t+1), u(t)\big) = f(\gamma_t\big(x_{rsd}(t : t + d)\big), \gamma_w(t : t + d))$$

[00054] Here f can be defined to be the pointwise sum of squares error $f = \sum_{i=t}^{i=t+d}(\gamma_t - \gamma_w)^2$ or a correlation function over the sequence such as the cosine similarity:

$$\frac{\gamma_t(x_{rsd}(t : t + d))^T \gamma_w(t : t + d)}{\|\gamma_t(x_{rsd}(t : t + d)\| \|\gamma_w(t : t + d)\|}$$

[00055] To test the performance of the algorithm, a time domain simulation of the drilling behavior is performed. Here a 2D drilling model based on the Pernedar Detournay (PD) model (L. Pernedar and E. Detournay, A Three-Dimensional Mathematical Model of Directional Drilling. PhD Thesis, 2013) is used to model the borehole propagation.

$$\chi \prod \Theta_{inc}(t) = -M_b[\Theta_{inc}(t) - < \Theta_{inc} >] + F_b[\Theta_{inc}(t) - \Theta_{inc}]$$
$$+ \sum_{i=1}^{n-1}\left[\frac{F_b M_i - F_i M_b - M_i \eta \prod}{\eta \prod}\right](< \Theta_{inc,i} > -$$
$$< \Theta_{inc} - \frac{\chi}{\eta} F_i \left(\frac{\Theta_{inc,i-1} - \Theta_{inc,i}}{x_i} - \frac{\Theta_{inc,i} - \Theta_{inc,i+1}}{x_{i+1}}\right)$$
$$+ \frac{F_b M_w - F_w M_b - M_w \eta \prod}{\eta \prod} \Upsilon sin < \Theta_i > - \frac{\chi}{\eta} F_w \Upsilon [\Theta_{inc}(t) - < \Theta_{inc} >] cos$$
$$< \Theta_i > \frac{F_b M_r - F_r M_b - M_r \eta \prod}{\eta \prod} (\Gamma_2)$$

[00056] Here M_i and F_i are coefficients related to the forces and moments of the drilling assembly. The coefficients ξ and η are related to the geometric design of the bit. $\Theta_{inc,i}$ is the inclination angle of the i stabilizer. $< \Theta_{inc,i} >$ is the inclination angle between the i stabilizer and the (i+1) stabilizer. $\prod$ is the weight-on-bit and Γ_2 is the control action applied. The position of the wellbore is obtained by integrating the inclination angle to give the x_{tvd} of the well.

$$x_{tvd} = \int_{t_0}^{t=T} cos\Theta_{inc} dt$$

[00057] The earth is modeled as a 2D unfaulted layer cake with parallel formation boundaries. The formation top is modeled as a gaussian process

$$x_f \sim GP(m(x), k(x, x'))$$

with a mean function drawn from a random trajectory, and an appropriate choice of a kernel k. The relative stratigraphic distance $x_{rsd}(t)$ is given by the difference of these two by

$$x_{rsd}(t) = x_f(t) - x_{tvd}$$

[00058] The petrophysical sensors with lower depth of investigation are given by a 1D table look up from a predetermined reference well $\gamma_t(x_{rsd})$.

[00059] Fig. 3 illustrates another embodiment of a work flow of the present invention where the trained reinforcement learning agent is a trained MCTS reinforcement learning agent. The Monte Carlo Tree Search Algorithm as described in Rémi Coulom (2007). "Efficient Selectivity and Backup Operators in Monte-Carlo Tree Search". Computers and Games, 5th International Conference, CG 2006, Turin, Italy, May 29–31, 2006. Revised Papers. H. Jaap van den Herik, Paolo Ciancarini, H. H. L. M. Donkers (eds.). is modified here to approximate a search over an enumerated set of all trajectories. These trajectories are captured by a tree data structure where the note represents stratigraphic position, and paths in the tree are feasible trajectories. The modified algorithm is illustrated in Fig. 3 as Monte Carlo Trajectory Sampling 200, where a distribution over trajectories is learned to guide the most probable trajectories for efficient searching.

[00060] For MCTS, inputs are:

- a system model A and
- a cost function that considers S steps of future cost,
- interpolation $I(\text{typelog RSD}, \text{typelog GR})$,
- initial formation dip D, and

- start RSD $rsd0$.

[00061] Sensor data INC_t, GR_t of length S are observed. Changes in RSD Δrsd over an advancement in trajectory length of length S are determined according to inclination sensor data INC_t and initial formation dip D. Based on the Δrsd, a Gaussian distribution $N(\mu, \sigma)$ is computed for actions. A simulation is performed using actions sampled out from $N(\mu, \sigma)$. Cost R of each trajectory sampled is computed according to GR (Gamma Ray) sensor data GR_t and the type log. The trajectory with least cost R has the best RSD sequence. The number of trajectories sampled is high when Δrsd is high, or vice versa.

[00062] Output of the MCTS is an RSD sequence that minimizes the cost function. The algorithm is preferably:

Initialize the current RSD, $curr_rsd = rsd0$

Initialize centre, $c = rsd0$

for t = 0,1, ... **do** */* Iterate over stands of length S */*

  Observe INC_t and GR_t */* Sensor data of length S */*

  Compute change in RSD Δrsd according to INC_t and initial formation dip D

  Update c according to Δrsd

  Compute mean $\mu = (c - curr_rsd)/S$

  Compute sigma σ according to μ

  / Search over N_p paths */*

  for p = 0,1, ... N_p **do**

    Set $\overline{rsd_0} = curr_rsd$ */* each path starts with curr_rsd */*

    Set $R(p) = 0$

    / estimating over stand length S */*

    for k = 0,1, ... S **do**

      Sample action a from Gaussian distribution $N(\mu, \sigma)$

      Predict the next RSD, $\overline{rsd_{k+1}} = A(\overline{rsd_k}, a)$

      Compute the cost r according to $GR(k), I(\overline{rsd_{k+1}})$

      Update $R(p) = \gamma R(p) + r$

    end

  end

  Select the best path with lower cost R

  Set $curr_rsd$ to the last RSD in best path

  Take RSD sequence in best path

end

[00063] Preferably, the reinforcement learning agent is trained using a simulation environment, more preferably using a simulation environment produced in accordance with

the method described in "Method for Simulating a Coupled Geological and Drilling Environment" filed in the USPTO on the same day as the present application, as provisional application US62/712490 filed 31 July 2018, the entirety of which is incorporated by reference herein.

[00064] For example, the reinforcement learning agent may be trained by (a) providing a training earth model defining boundaries between formation layers and petrophysical properties of the formation layers in a subterranean formation comprising data selected from the group consisting of seismic data, data from an offset well and combinations thereof, and producing a set of model coefficients; (b) providing a toolface input corresponding to the set of model coefficients to a drilling attitude model for determining a drilling attitude state; (c) determining a drill bit position in the subterranean formation from the drilling attitude state; (d) feeding the drill bit position to the training earth model, and determining an updated set of model coefficients for a predetermined interval and a set of signals representing physical properties of the subterranean formation for the drill bit position; (e) inputting the set of signals to a sensor model for producing at least one sensor output and determining a sensor reward from the at least one sensor output;(f) correlating the toolface input and the corresponding drilling attitude state, drill bit position, set of model coefficients, and the at least one sensor output and sensor reward in the simulation environment; and (g) repeating steps b) – f) using the updated set of model coefficients from step d).

[00065] The drilling model for the simulation environment may be a kinematic model, a dynamical system model, a finite element model, and combinations thereof.

EXAMPLES 1 – 4

[00066] The method of the present invention was tested. Referring now to Figs. 4 – 7, a synthetic well was generated based on an actual gamma ray log. The real data is identified by a type log gamma ray plot 62. Based on the type log gamma ray plot 62, a boundary 64 representing the top of a target formation was determined and a synthetic true well path 66 was generated. Region 72 represents a 1.5-m (5-foot) error about the true well path 66, while region 74 represents a 3-m (10-foot) error about the well path 66. The goal of the test was to match the true well path 66 as best as possible.

[00067] In each of Example 1 – 4, a Bayesian reinforcement learning agent was trained according to the method described in co-pending application entitled "Method for Simulating

a Coupled Geological and Drilling Environment" filed in the USPTO on the same day as the present application.

[00068] Well log gamma ray data 76 was fed to the trained agent and a set of control inputs, in this case well inclination angle 78, was used to steer the well-boring along the true well path 66, according to the method described in co-pending application entitled "Process for Training a Deep Learning Process for Geological Steering Control" filed in the USPTO on the same day as the present application, as provisional application US62/712506 filed 31 July 2018, the entirety of which is incorporated by reference herein.

[00069] The well path 82 resulting from the Bayesian reinforcement learning agent and the well path 84 resulting from the trained agent with mean square error demonstrated good fit to the true well path 66. As shown in Figs. 4 – 7, the fit of well paths 82 and 84 improved over time with a reward function described in the autonomous geosteering method.

[00070] While preferred embodiments of the present disclosure have been described, it should be understood that various changes, adaptations and modifications can be made therein without departing from the spirit of the invention(s) as claimed below.

CLAIMS

1. A method of geosteering in a wellbore construction process, the method comprising the steps of:
   - providing an earth model defining boundaries between formation layers and petrophysical properties of the formation layers in a subterranean formation comprising data selected from the group consisting of seismic data, data from an offset well and combinations thereof;
   - comparing sensor measurements related to the wellbore construction process to the earth model;
   - obtaining an estimate from the earth model for a relative geometrical and geological placement of the well path with respect to a geological objective using a trained reinforcement learning agent; and
   - determining an output action based on the sensor measurement for influencing a future profile of the well path with respect to the estimate.
2. The method of claim 1, wherein the trained reinforcement learning agent is a trained Bayesian reinforcement learning agent.
3. The method of claim 1, wherein the trained reinforcement learning agent is a trained Monte Carlo Trajectory Sampling reinforcement learning agent.
4. The method of claim 1, wherein the output action is determined by maximizing the placement of the well path with respect to a geological datum.
5. The method of claim 4, wherein the geological datum is selected from the group consisting of a rock formation boundary, a geological feature, an offset well, an oil/water contact, an oil/gas contact, an oil/tar contact and combinations thereof.
6. The method of claim 4, wherein the estimate is determined by providing to the trained reinforcement learning agent:
   - a state space representation for a given depth for a position and a direction of the well path and the geological datum, having a discretized representation of the output action as a set of plausible geological datum changes;

- a state transition function for determining a transition between the state space representation at depth t and depth t+1 conditional upon the output action;
- an observational model for modeling the sensor measurements to the earth model;
- a reward function;
- a discount rate applied to the reward function for determining a discounted reward function; and
- a value function representing a past sum of discounted rewards for the transition of depth running forward in time.

7. The method of claim 6, wherein an optimal output action for a most probable well path is solved with respect to the value function to minimize or maximize the expected sum of the reward function at a given depth.

8. The method of claim 6, wherein an optimum value function is determined by iterating on a maximum or minimum of the expected sum of the reward function at depth t with the value of the state space at depth t-1 with respect to state transition function, selecting the highest value state with respect to a constraint, and propagating forward in depth the output actions to determine an optimum formation interpretation.

9. The method of claim 6, wherein the state space is continuous.

10. The method of claim 6, the state transition function is pretrained on historical wells and or synthetic data, wherein the pretraining is selected from the group consisting of a neural network, a probabilistic graphical model, and combinations thereof.

11. The method of claim 6, wherein the discounted sum of rewards is based on discretized depth intervals in an arc length of the well path.

12. The method of claim 6, wherein the reward function is selected from the group consisting of a sequence similarity measure, a mean squared error reward function, a Huber loss reward function, a non-convex reward function and combination thereof.

13. The method of claim 6, wherein the observation model is a look-up from a type log or the earth model.

14. The method of claim 1, wherein the earth model is a static model.

15. The method of claim 1, wherein the earth model is a dynamic model that changes dynamically during the drilling process.

16. The method of claim 1, wherein the sensor measurements are provided as a streaming sequence.

17. The method of claim 1, wherein the sensor measurements are measurements obtained from sensors selected from the group consisting of gamma-ray detectors, neutron density sensors, porosity sensors, sonic compressional slowness sensors, resistivity sensors, nuclear magnetic resonance, mechanical properties, inclination, azimuth, roll angles, and combinations thereof.

18. The method of claim 1, wherein the reinforcement learning agent is trained in a simulation environment.

19. The method of claim 18, wherein the simulation environment is produced by a training method comprising the steps of:

a) providing a training earth model defining boundaries between formation layers and petrophysical properties of the formation layers in a subterranean formation comprising data selected from the group consisting of seismic data, data from an offset well and combinations thereof, and producing a set of model coefficients;

b) providing a toolface input corresponding to the set of model coefficients to a drilling attitude model for determining a drilling attitude state;

c) determining a drill bit position in the subterranean formation from the drilling attitude state;

d) feeding the drill bit position to the training earth model, and determining an updated set of model coefficients for a predetermined interval and a set of signals representing physical properties of the subterranean formation for the drill bit position;

e) inputting the set of signals to a sensor model for producing at least one sensor output and determining a sensor reward from the at least one sensor output;

f) correlating the toolface input and the corresponding drilling attitude state, drill bit position, set of model coefficients, and the at least one sensor output and sensor reward in the simulation environment; and

g) repeating steps b) – f) using the updated set of model coefficients from step d).

20. The method of claim 19, wherein the drilling attitude model is selected from the group consisting of a kinematic model, a dynamical system model, a finite element model, and combinations thereof.

21. The method of claim 1, wherein the output action is selected from the group consisting of curvature, roll angle, set points for inclination, set points for azimuth, Euler angle, rotation matrix quaternions, angle axis, position vector, position Cartesian, polar, and combinations thereof.

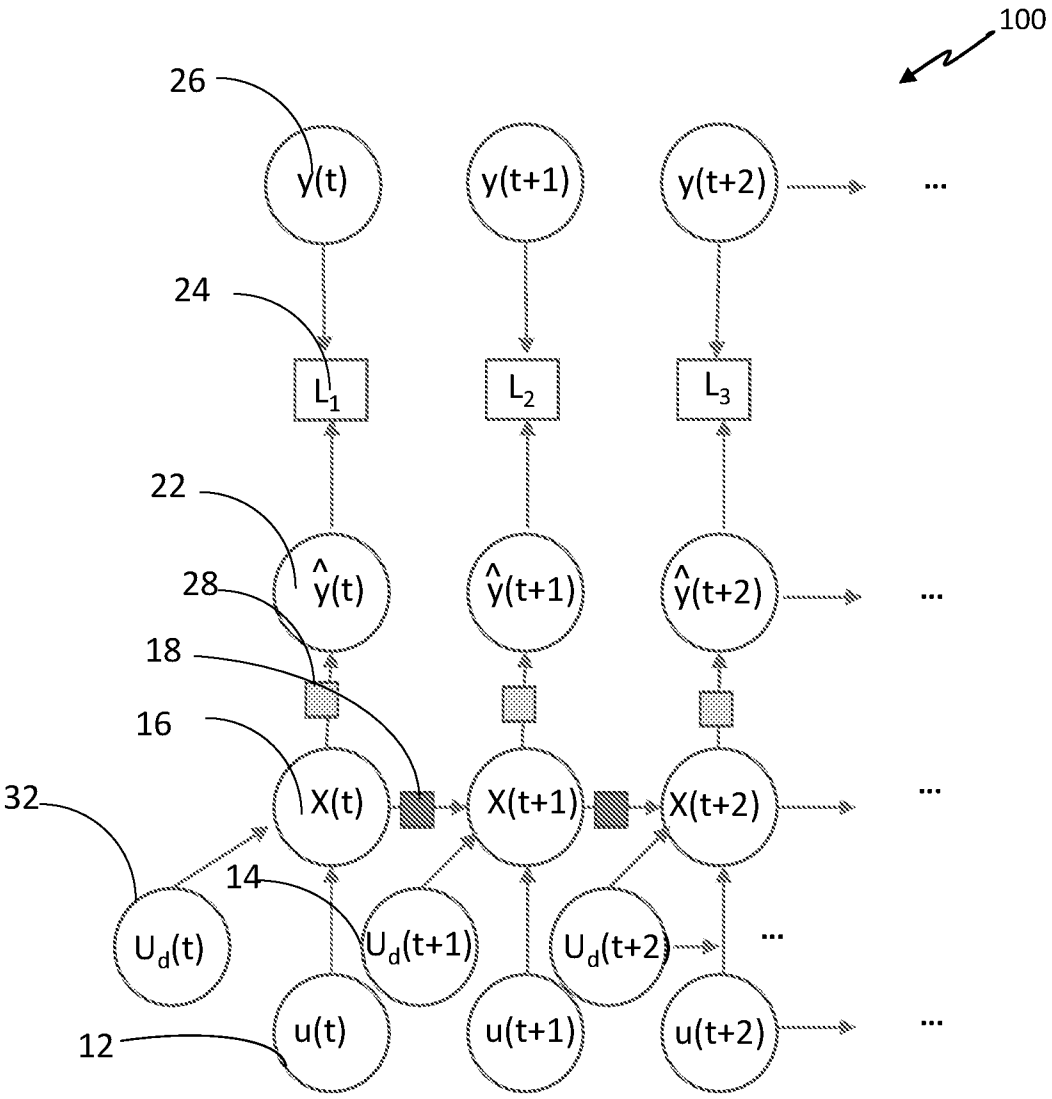

100
26
y(t)
y(t+1)
y(t+2)
...
24
L1
L2
L3
22
ŷ(t)
ŷ(t+1)
ŷ(t+2)
28
18
16
X(t)
X(t+1)
X(t+2)
32
14
Ud(t)
Ud(t+1)
Ud(t+2)
12
u(t)
u(t+1)
u(t+2)

FIG. 1

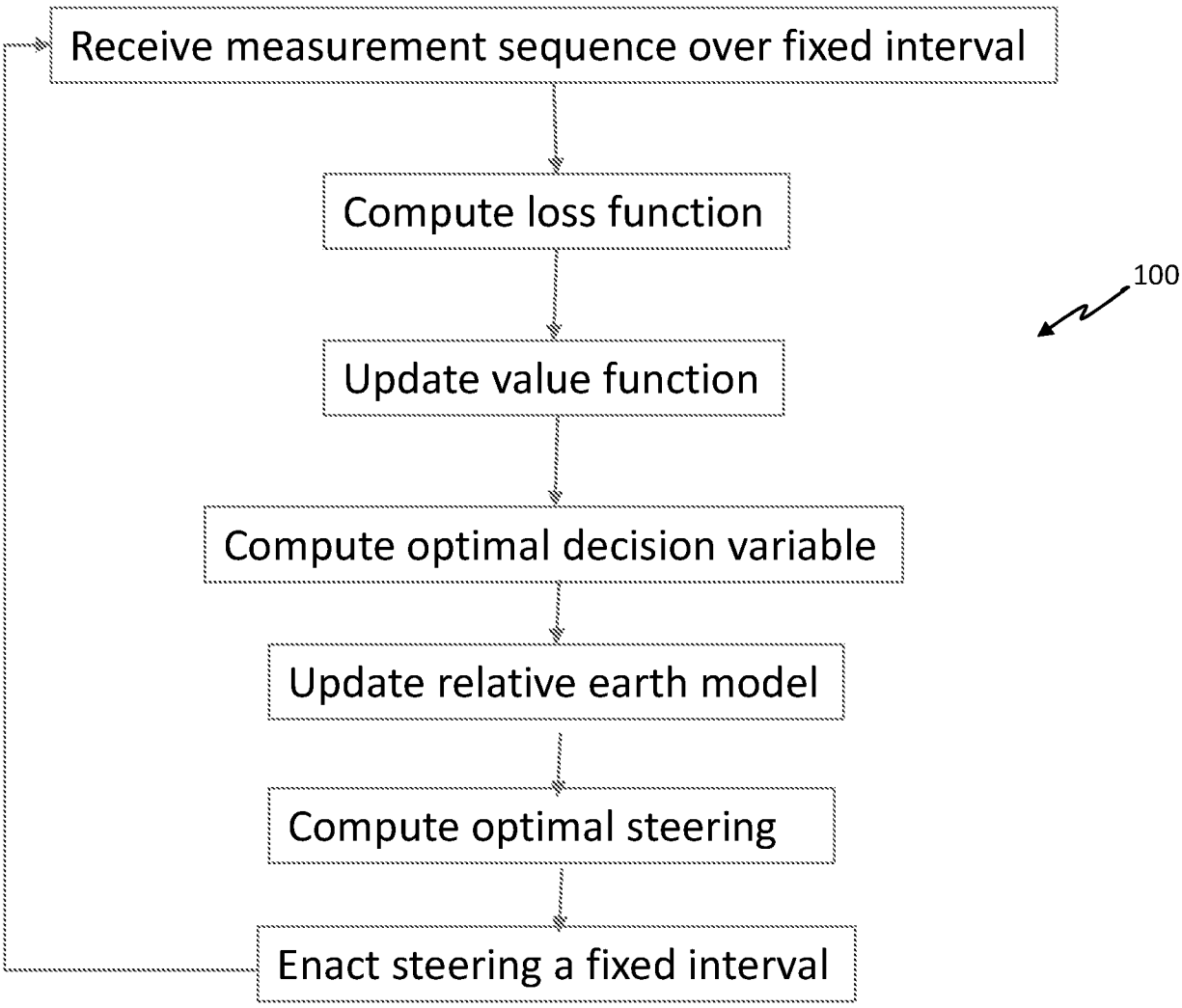

Receive measurement sequence over fixed interval
Compute loss function
Update value function
Compute optimal decision variable
Update relative earth model
Compute optimal steering
Enact steering a fixed interval
100

FIG. 2

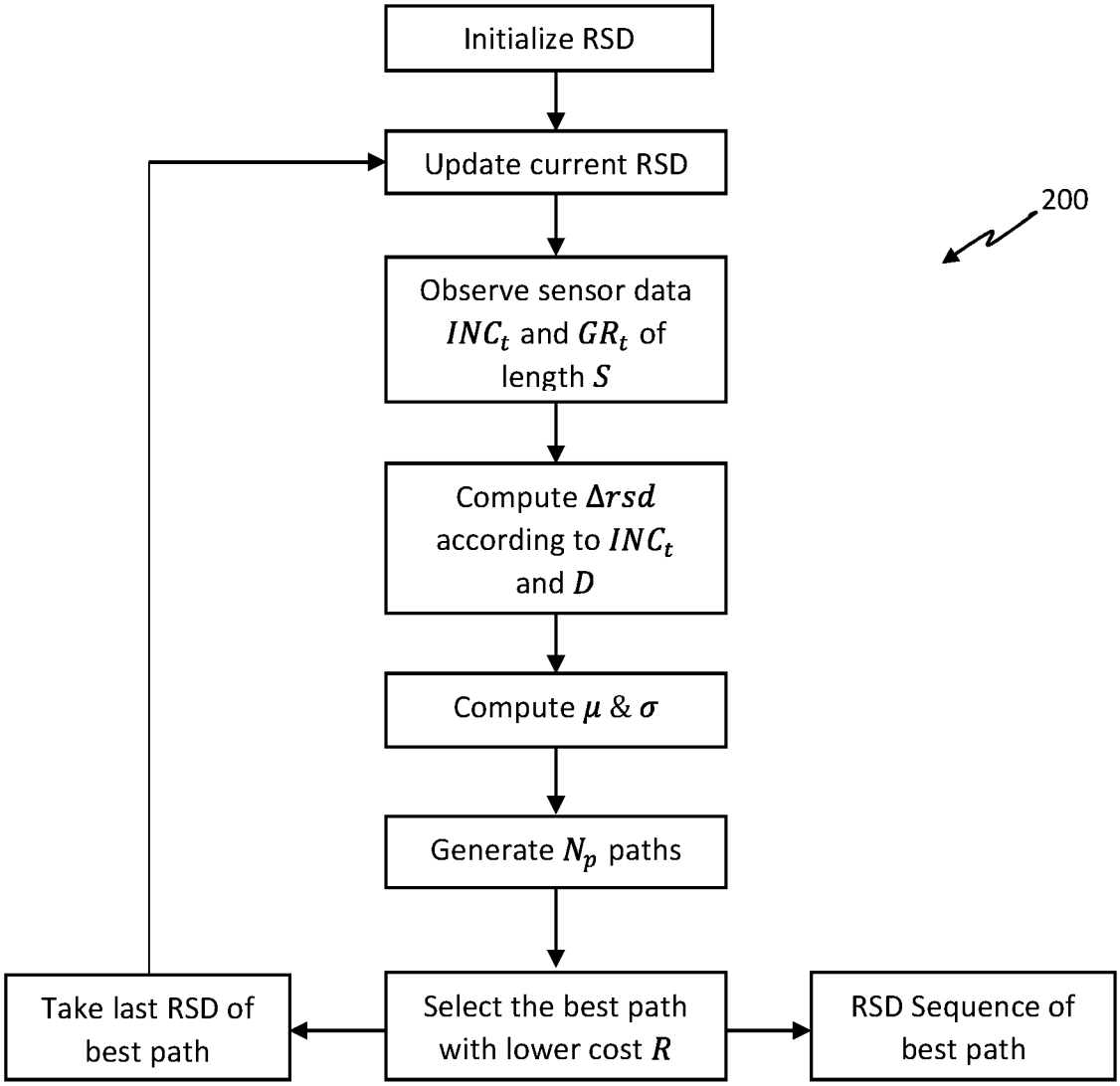

Initialize RSD
Update current RSD
200
Observe sensor data INC_t and GR_t of length S
Compute Δrsd according to INC_t and D
Compute μ & σ
Generate N_p paths
Take last RSD of best path
Select the best path with lower cost R
RSD Sequence of best path

FIG. 3

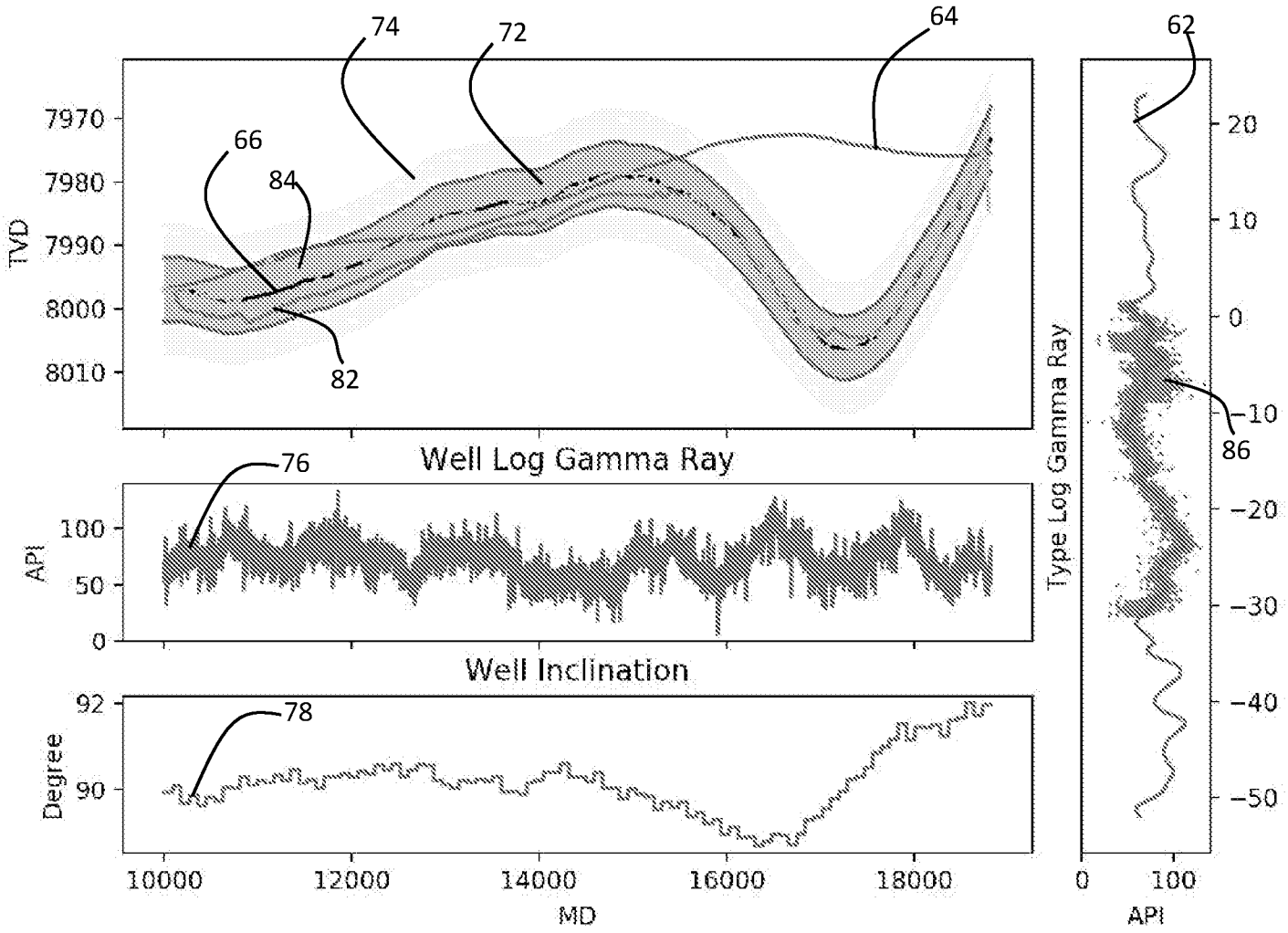

74
72
64
62
66
84
82
86
76
78
TVD
7970
7980
7990
8000
8010
Well Log Gamma Ray
API
100
50
0
Well Inclination
Degree
92
90
10000
12000
14000
16000
18000
MD
Type Log Gamma Ray
20
10
0
−10
−20
−30
−40
−50
0
100
API

FIG. 4

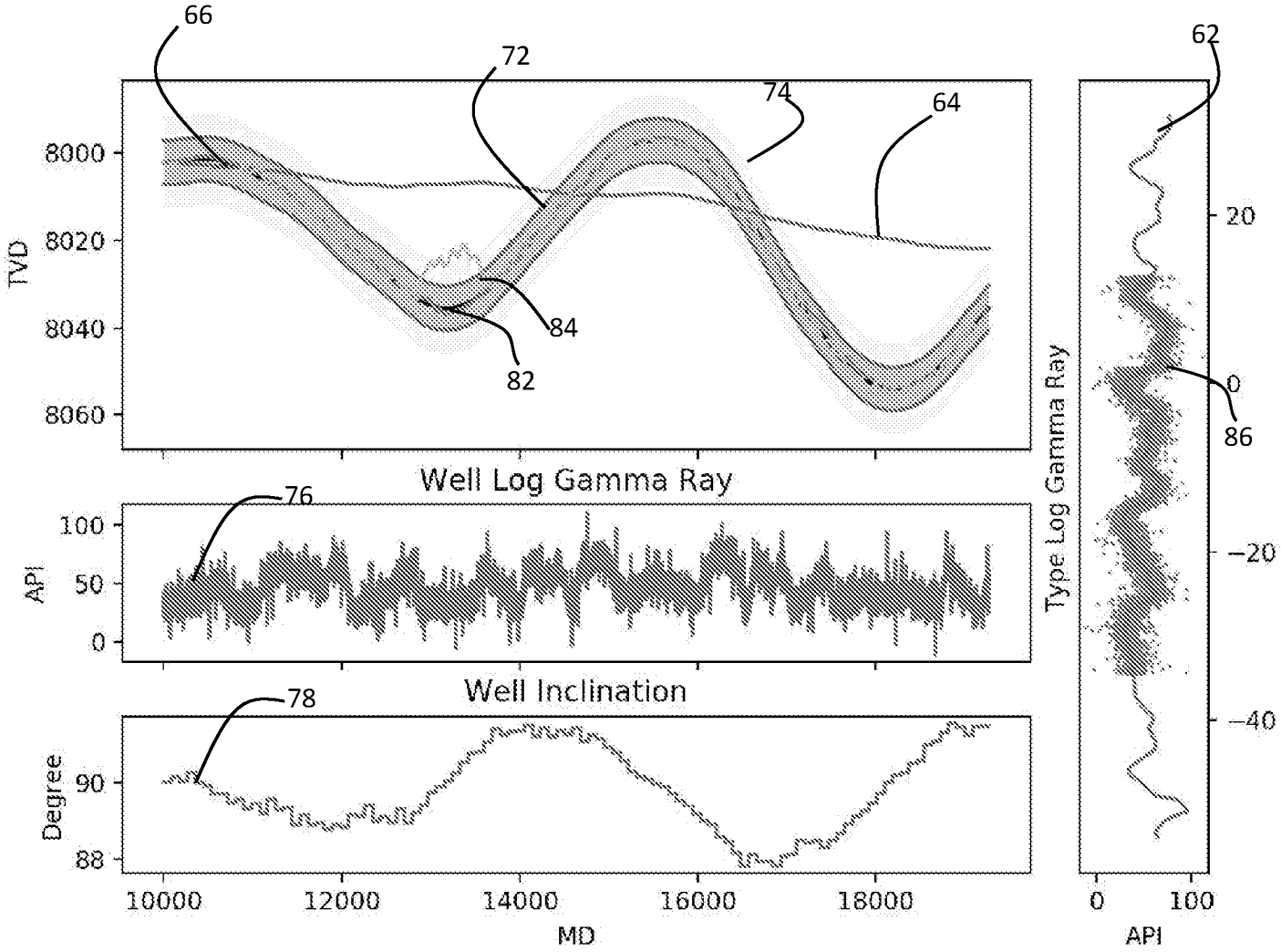

66
72
74
64
62
8000
8020
8040
8060
TVD
84
82
Well Log Gamma Ray
76
100
50
0
API
Type Log Gamma Ray
20
0
−20
−40
86
Well Inclination
78
Degree
90
88
10000
12000
14000
16000
18000
MD
0
100
API

FIG. 5

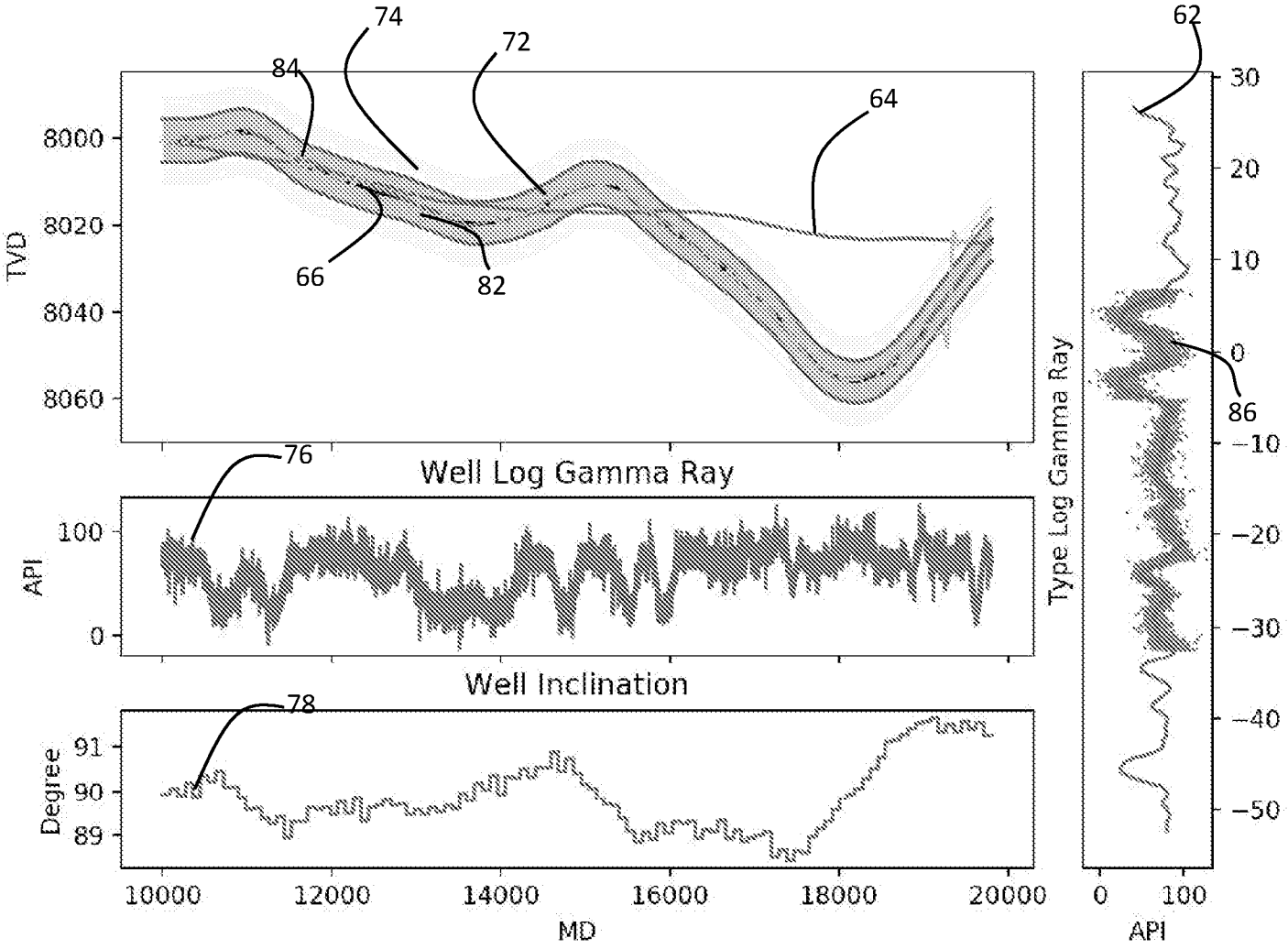

74
72
84
64
62
66
82
86
76
78
TVD
8000
8020
8040
8060
Well Log Gamma Ray
API
100
0
Well Inclination
Degree
91
90
89
10000
12000
14000
16000
18000
20000
MD
Type Log Gamma Ray
30
20
10
0
−10
−20
−30
−40
−50
0
100
API

FIG. 6

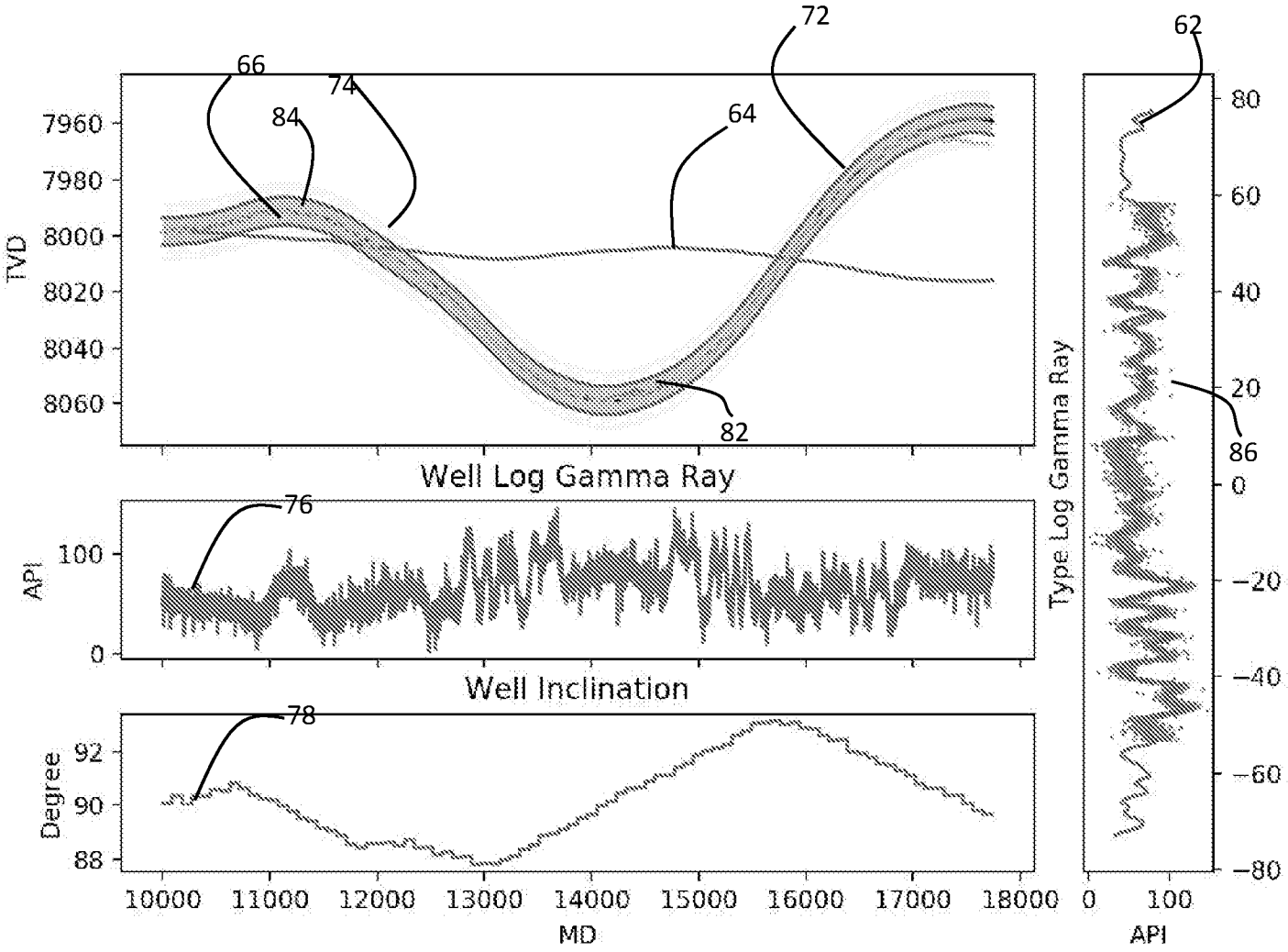

72
62
66
74
84
64
7960
7980
8000
8020
8040
8060
TVD
82
Well Log Gamma Ray
76
API
100
0
Type Log Gamma Ray
80
60
40
20
86
0
−20
−40
−60
−80
Well Inclination
78
Degree
92
90
88
10000
11000
12000
13000
14000
15000
16000
17000
18000
MD
0
100
API

FIG. 7

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/044042

A. CLASSIFICATION OF SUBJECT MATTER
INV. E21B7/04 G06Q50/02
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)
E21B G06Q

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, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2018/106748 A1 (SCHLUMBERGER TECH CORPORATION [US]; SCHLUMBERGER CA LTD [CA]; SERVICES) 14 June 2018 (2018-06-14)	1,3-21
Y	figures 1-4, 9, 14 paragraphs [0005], [0027], [0028], [0037], [0050], [0054] - [0058], [0068] - [0073] paragraphs [0083] - [0084], [0087] - [0088], [0093], [0105] - [0109], [0141] - [0148], [0155] - [0157] paragraphs [0180], [0187] - [0188], [0194]	2

	-/--	

[X] Further documents are listed in the continuation of Box C.

[X] 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	Date of mailing of the international search report
19 November 2019	25/11/2019

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Pasquini, Matteo

3

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/044042

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2017/188858 A1 (SOFRONOV IVAN L'VOVICH [RU]; SCHLUMBERGER CA LTD [CA] ET AL.) 2 November 2017 (2017-11-02) paragraphs [0025] - [0029], [0068] - [0072], [0079] - [0083], [0094], [0127], [0134], [0156]; figures 1-4,6-15	1,4-21
Y	MOHAMMED GHAVAMZADEH ET AL: "Bayesian Reinforcement Learning: A Survey", FOUNDATIONS AND TRENDS IN MACHINE LEARNING, vol. 8, no. 5-6, 1 January 2015 (2015-01-01), pages 359-492, XP055542023, ISSN: 1935-8237, DOI: 10.1561/2200000049	2
A	abstract Paragraphs: 2.5, 4.5, 5, 7 Appendix: A	1,3-21
A	Jacob Pollock ET AL: "Machine Learning for Improved Directional Drilling", , 30 April 2018 (2018-04-30), XP055542591, DOI: 10.4043/28633-MS Retrieved from the Internet: URL:https://onepetro.org/download/conference-paper/OTC-28633-MS?id=conference-paper/OTC-28633-MS [retrieved on 2019-01-15] the whole document	1,4-15
A	HUGH WINKLER: "Geosteering by exact inference on a Bayesian network", GEOPHYSICS, vol. 82, no. 5, 20 July 2017 (2017-07-20), pages D279-D291, XP055565580, US ISSN: 0016-8033, DOI: 10.1190/geo2016-0569.1 page 279, page 284: "results", caption of fig.4, and page 287 par.2-3	1-21
A	R De Togni: "PREDICTIVE MODEL FOR DRILLING PHASE DURATION OF OIL & GAS WELLS", , 25 July 2018 (2018-07-25), XP055637737, Retrieved from the Internet: URL:https://www.politesi.polimi.it/bitstream/10589/141802/3/Tesi_De_Togni.pdf [retrieved on 2019-10-30] par.1.1, 1.2, 3.4.4, 4.2.2, 4.2.3, 5.1 and 6.2	1-21

-/--

3

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/044042

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2016/186553 A1 (MULLINS OLIVER C [US] ET AL) 30 June 2016 (2016-06-30) The whole document and in paticular fig.4 and par.3 -----	1,4,5

3

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2019/044042

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2018106748 A1	14-06-2018	EP 3552047 A1	16-10-2019
		US 2019302310 A1	03-10-2019
		WO 2018106748 A1	14-06-2018
WO 2017188858 A1	02-11-2017	NONE	
US 2016186553 A1	30-06-2016	CA 2725133 A1	26-11-2009
		CN 102037212 A	27-04-2011
		US 2011061935 A1	17-03-2011
		US 2015008037 A1	08-01-2015
		US 2016186553 A1	30-06-2016
		WO 2009142868 A2	26-11-2009